University Libraries
University of Nevada, Las Vegas

1-1-2020

Wind Power Forecasting Methods Based on Deep Learning: A Survey

Xing Deng
Jiangsu University of Science and Technology

Haijian Shao
Jiangsu University of Science and Technology, jsj_shj@just.edu.cn

Chunlong Hu
Jiangsu University of Science and Technology

Dengbiao Jiang
Jiangsu University of Science and Technology

Yingtao Jiang
University of Nevada, Las Vegas, yingtao.jiang@unlv.edu
Follow this and additional works at: https://digitalscholarship.unlv.edu/ece_fac_articles

Part of the Agriculture Commons, Artificial Intelligence and Robotics Commons, and the Power and Energy Commons

Wind Power Forecasting Methods Based on Deep Learning: A Survey

Xing Deng[1, 2], Haijian Shao[1, *], Chunlong Hu[1], Dengbiao Jiang[1] and Yingtao Jiang[3]

Abstract: Accurate wind power forecasting in wind farm can effectively reduce the enormous impact on grid operation safety when high permeability intermittent power supply is connected to the power grid. Aiming to provide reference strategies for relevant researchers as well as practical applications, this paper attempts to provide the literature investigation and methods analysis of deep learning, enforcement learning and transfer learning in wind speed and wind power forecasting modeling. Usually, wind speed and wind power forecasting around a wind farm requires the calculation of the next moment of the definite state, which is usually achieved based on the state of the atmosphere that encompasses nearby atmospheric pressure, temperature, roughness, and obstacles. As an effective method of high-dimensional feature extraction, deep neural network can theoretically deal with arbitrary nonlinear transformation through proper structural design, such as adding noise to outputs, evolutionary learning used to optimize hidden layer weights, optimize the objective function so as to save information that can improve the output accuracy while filter out the irrelevant or less affected information for forecasting. The establishment of high-precision wind speed and wind power forecasting models is always a challenge due to the randomness, instantaneity and seasonal characteristics.

Keywords: Deep learning, reinforcement learning, transfer learning, wind power forecasting.

Notation
ANFIS: Adaptive neuro-fuzzy inference system;
CNNs: Convolution neural networks;
DBNs: Deep belief networks;
DNNs: Deep neural networks;
DQN: Reinforcement learning Q-network;

[1]School of Computer Science, Jiangsu University of Science and Technology, Zhenjiang, China.
[2]School of Automation, Key Laboratory of Measurement and Control for CSE, Ministry of Education, Southeast University, Nanjing, China.
[3]Department of Electrical and Computer Engineering, University of Nevada, Las Vegas, USA.
*Corresponding Author: Haijian Shao. Email: jsj_shj@just.edu.cn.

DRL: Combination-deep reinforcement learning;
ELM: Extreme learning machine;
EO: Extremal Optimization;
EMD: Empirical model decomposition;
GCDLA: Graph convolutional deep learning architecture;
GRU: Gated recurrent unit;
KPCA: Kernel Principal Component Analysis;
LSTM: Long Short-Term Memory;
MAPE: Absolute percentage errors;
MDN: Ensemble mixture density network;
MPPs: Maximum power points;
MPPT: Maximum power point tracking training;
NP: Non-deterministic Polynomial hard;
NWP: Numerical weather prediction;
NREL: National Renewable Energy Laboratory;
PSO: particle swarm optimization;
PSR: Phase Space Reconstruction;
QR: Spine quantile regression;
VMD: Variational model decomposition;
WPD: Wavelet Packet Decomposition;
RL: Reinforcement learning;
RNNs: Recurrent neural networks;
SAE: Stacked auto-encoders;
SVR: Support vector regression.

1 Introduction

Wind Power has been recognized as a promising alternative, clean energy source to fossil fuel-generated electricity. Produced by capturing the air flow that occurs in the atmosphere, wind power by nature is found to be random and instantaneous in yield, and season-dependent in short-term. For safe operation of wind mill and distribution of electricity generated from wind power, it is important to accurately forecast wind speed. Unfortunately, accurate wind speed forecasting, especially short-term wind speed forecasting (STWSF), remain quite a challenge [Wang, Zhang, Xu et al. (2018); Shao, Deng and Jiang (2018); Zhang, Sun, Sun et al. (2016)]. At present, wind power prediction models can be coarsely divided into three categories: physics-based prediction model, statistics-based prediction model and hybrid prediction model.

(i) Physics-based prediction models. This kind of models take into account the weather phenomena (or weather processes) and treat weather changes as non-random events that satisfy certain physical laws, such as energy conservation [Hodge and Milligan (2011); Shao, Deng and Cui (2016); Wang, Liu, Yu et al. (2019)]. There is one fundamental assumption about the physics-based models that at any given moment, the weather condition (state) is determined by the atmospheric data of the past (history). That is, the atmospheric change can be simulated/calculated through the numerical weather prediction

(NWP) based on the air pressure, temperature, roughness, contour, roughness and obstacles around the wind farm. Physics-based models are able to predict the desired variables directly from real-time data, making them preferred models for wind farms. However, since this type of models requires real-time, high-precision data, there is a high demand for accurate data acquisition systems and high-speed data transmission networks. In addition, the modeling process tends to be complicated and sensitive to the errors introduced by the initial condition.

(ii) Statistics-based prediction models. This type of models begin with certain weather phenomena (or weather processes), and the evolving weather changes are treated as a random process. Unlike the previously described physics-based models, statistics-based models at different runs may predict different weather outcomes for the same ambient conditions, as the models rely on calculating the probability of occurrence of certain weather situations. As an improvement, this kind of models may sometimes include variables that have physical meanings and allow people to interpret what is going on with the weather. Quite popular among researchers and wind farm operators, statics-based models, however, place too much weight on the historical data, and they are still lack of physics to back up their prediction results [Shao and Deng (2018); Wu, Zhu, Su et al. (2015); Wang, Ji, Xue et al. (2016)]. Another drawback of these statics-based models is these models fail to consider the time delay in the forecasting process, and it becomes difficult to integrate the wind power time series into the forecast model.

(iii) Combined prediction models. This kind of models is weighted combination of different prediction models. These combined models hold the promise of avoiding the weakness exhibited by a single model in wind power forecasting. One popular combined prediction model uses wavelet transformation to reduce the negative effects of non-stationary and non-smooth wind power time series to improve prediction accuracy, after which neural network models can be applied for better prediction of winds. In a nutshell, a combined model is able to take into account the various features of the wind power time series, which are multi-time scale and multi-resolution in nature [Hong, Li, Geng et al. (2019); Pant, Han and Wang (2019); Zhang and Hong (2019)]. Combined prediction models benefit from the rapid development of artificial intelligence, including deep learning and reinforcement learning, to help improve the classification accuracy as well as reveal the complex nonlinear relationship in wind power forecasting. The advantages of deep learning in feature selection and function approximation compared with the traditional feedback neural networks guarantees the its superior processing ability in pattern recognition and computer vision etc. In particular, the transfer learning and reinforcement learning methods etc can be used to combine with the deep learning, and then facilitates the rapid improvement in feature analysis and calculating efficiency. This paper thus is dedicated to survey how deep learning, enforcement learning and transfer learning are applied in wind speed and wind power forecasting, especially in the areas of data processing of wind power, input features selection, forecasting model framework establishment and model structure optimization. Our goal here is to provide an assessment of various network structures including feed-forward network and Recurrent Neural Network (RNN), among a few others, and determine their strengths and applicability in the context of wind modeling

methods as described above.

The rest of this paper is organized as follows. In Section 2, we all introduce the fundamental forecasting framework, consisting of data preprocessing and feature selection and analysis. Section 3 is devoted to make comparisons of various studies reported in the literature concerning deep learning and reinforcement learning techniques, after which model structure optimization strategies based on the deep learning will be surveyed in Section 4. Finally, Section 5 concludes the paper with discussions on future research.
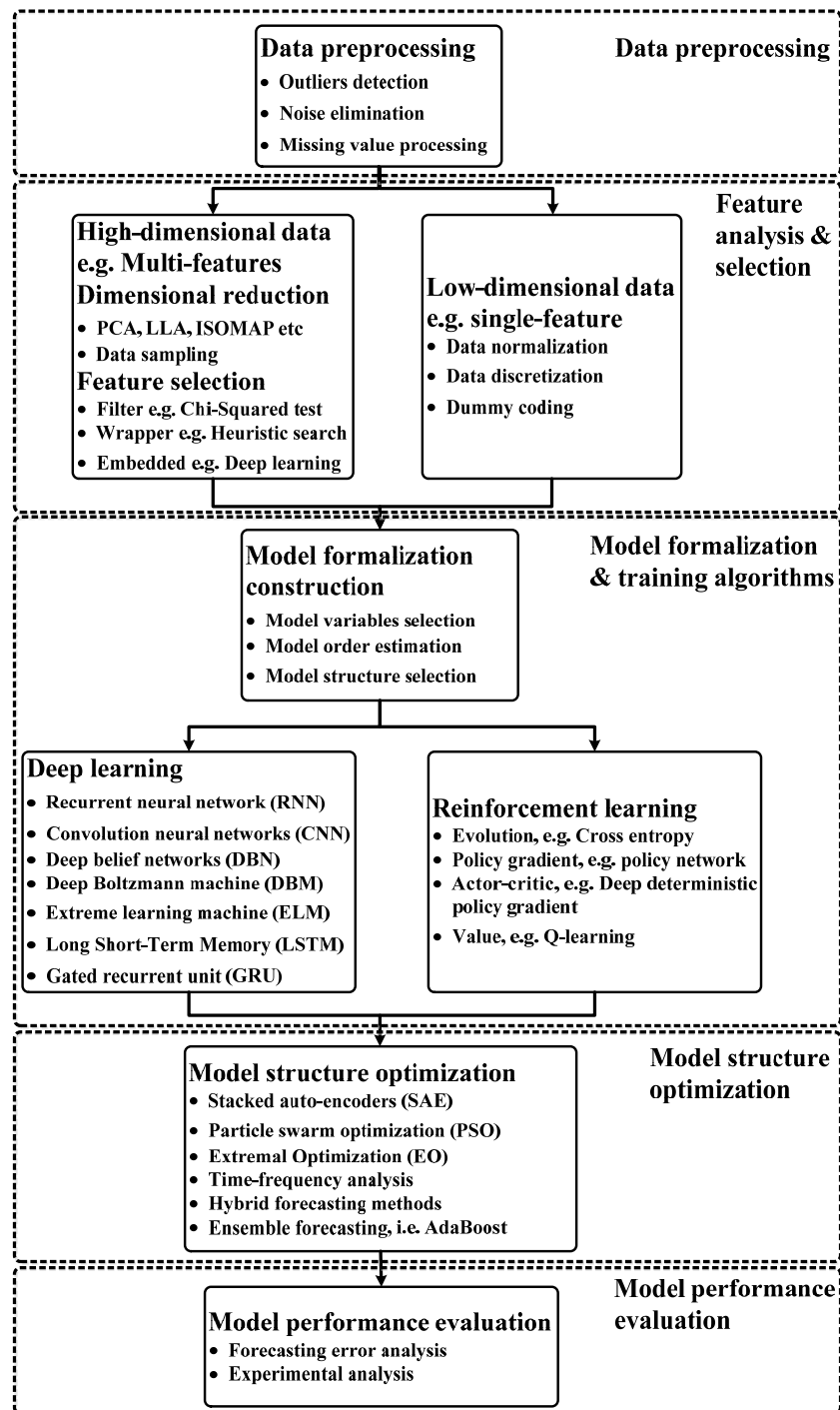
2 Forecasting frameworks

The general wind speed forecasting framework mainly includes five components: data preprocessing, high-dimensional and low-dimensional data feature analysis, model formalization based on deep learning and reinforcement learning, model structure optimization and model performance evaluation. All these components will be discussed in this section, the general forecasting framework is given in Fig. 1.

2.1 Data preprocessing

Normally, the data derived from the wind farm contains many outliers, noisy data and missing value caused by inevitable factors. Due to the continuity of data in time as well the true information covered by noise, these data should be identified correctly by appropriate preprocessing algorithms. Outlier detection, noise elimination and missing value filling are the three most common and important aspects in data preprocessing. Due to inconsistent data (usually represented by the same input but shown in different outputs, sometimes caused by human record error), irrelevant data and local extremum, resulting that the deviation of the forecasting model are caused. In particular, sparse data may be weakened in distribution. The method commonly used for local extremum in engineering is to normalize the data to a certain ratio while maintaining the specific structure, such as data discretization and normalization, which also are used to remove the unit limits for dimensionless and weighted processing. Assuming that all data are normally distributed, then the data is model-processed, and the actual data can be obtained if the reverse operation is performed. Missing values usually appeared in the actual collected data from wind farm, which destroys the samples' integrity over time continuity. Based on the available variables around the missing value and their relationship, the interpolation function is most widely used to complete the missing values of the wind farm. Noisy data are one of the most common problems in wind farm. The most common used method is to utilize the filter method to remove the noise and obtain the actual signal estimation. The most ideal filtering technique reflecting at the corrupted data removing, effectively clean up noise while minimize signal distortion. According to the current filtering method, the commonly used wind farm data filtering strategies are divided into three categories:

(i) Time-domain filtering. The median or mean of the measurement data is placed in a window with a predetermined size, and then the threshold is set in the window to determine whether the current data is reasonable. The more robust and reasonable data will be calculated and used to replace the unreasonable ones. Since this method is always obtained

Data preprocessing
- Outliers detection
- Noise elimination
- Missing value processing

Data preprocessing

High-dimensional data e.g. Multi-features Dimensional reduction
- PCA, LLA, ISOMAP etc
- Data sampling

Feature selection
- Filter e.g. Chi-Squared test
- Wrapper e.g. Heuristic search
- Embedded e.g. Deep learning

Low-dimensional data e.g. single-feature
- Data normalization
- Data discretization
- Dummy coding

Feature analysis & selection

Model formalization construction
- Model variables selection
- Model order estimation
- Model structure selection

Model formalization & training algorithms

Deep learning
- Recurrent neural network (RNN)
- Convolution neural networks (CNN)
- Deep belief networks (DBN)
- Deep Boltzmann machine (DBM)
- Extreme learning machine (ELM)
- Long Short-Term Memory (LSTM)
- Gated recurrent unit (GRU)

Reinforcement learning
- Evolution, e.g. Cross entropy
- Policy gradient, e.g. policy network
- Actor-critic, e.g. Deep deterministic policy gradient
- Value, e.g. Q-learning

Model structure optimization
- Stacked auto-encoders (SAE)
- Particle swarm optimization (PSO)
- Extremal Optimization (EO)
- Time-frequency analysis
- Hybrid forecasting methods
- Ensemble forecasting, i.e. AdaBoost

Model structure optimization

Model performance evaluation
- Forecasting error analysis
- Experimental analysis

Model performance evaluation

Figure 1: The general wind power forecasting framework

through the absolute values, the corresponding estimation interval with upper and lower bounds can be directly obtained. If the data can be fitted by a normal distribution within a certain time interval, then the corresponding confidence interval estimate can also be calculated. The median filter as one of the most used time-domain filter methods used for the wind speed preprocessing are shown in Fig. 2.
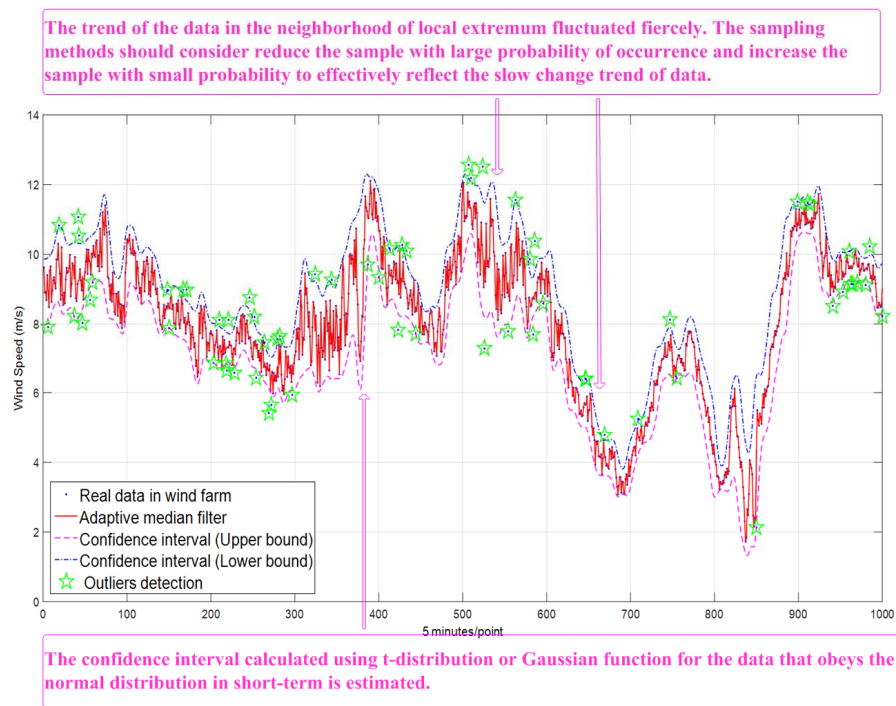
(ii) Frequency-domain filtering. The data is decomposed into low-frequency and high-frequency components by applying Fourier transform; the data's high-frequency component is then filtered out, so that the data are smoothed out. The only drawback of the Fourier transform is that one cannot simultaneously perform time-frequency domain analysis at the same time. Time-frequency domain filtering can also be obtained by multiplying the signal in the frequency domain by a piecewise function or by convolving with the Sinc function in the time domain.

(iii) Time-frequency domain filtering. The measured data is simultaneously transformed at the time and the frequency domains to obtain effective features of signals over a wide range. The premise of data filtering is that it assumes that data has sufficient knowledge of its neighborhood and thus filtering causes no or litter loss of the useful information. Recognized as the most widely used methods, wavelet analysis as an ideal time-frequency domain local analysis method where its time domain and frequency domain windows can be changed. For the low frequency component of the signal, the high frequency resolution and low time resolution of the time series can be analyzed while higher time resolution and lower frequency resolution can be obtained in high frequency component. Especially at large scales, the low-frequency global information of the signal can be represented, and the high-frequency local features can be represented at a small scale. Filtering with wavelet transforms is one of the most reprehensive used time-frequency methods in noisy data processing, and has been shown in Fig. 3. In addition, adaptive filtering methods such as empirical mode decomposition (EMD) [Wang, Zhang, Li et al. (2014); Liu, Chen, Tian et al. (2012); Wang, Zhang, Wu et al. (2016); Ren, Suganthan and Srikanth (2014a,b)], variational mode decomposition (VMD) Lahmiri (2015); Yang, Peng, Wei et al. (2016); Wang, Luo, Grunder et al. (2017); Li, Jiang, Guo et al. (2018); Naik, Dash, Dash et al. (2018); Zhang, Liu, Qin et al. (2016)], ensemble empirical mode decomposition (EEMD) [Sun and Wang (2018); An, Jiang, Li et al. (2011); Feng, Liang, Zhang et al. (2012); Yeh, Shieh and Huang (2010); Wang, Zhang, Wu et al. (2016)], is based on the linear and steady-state spectrum analysis method related to Fourier transformation, which can utilize various of time scale for signal decomposition without setting any base function, to reduce the effect of the non-smoothness of wind power data on forecasting modeling.

2.2 Feature analysis and selection

The structural design of nonlinear multi-inputs learning system is always a challenge due to the difficulties in feature selection which is of fundamental importance. Defining a set of optimized input variables or features should first take into consideration of the effects of the inputs variables on the forecasting model. Some useful and most widely used criteria are given as follows,

(i) Correlation analysis. Usually, operators or engineers usually focus that if the number

The trend of the data in the neighborhood of local extremum fluctuated fiercely. The sampling methods should consider reduce the sample with large probability of occurrence and increase the sample with small probability to effectively reflect the slow change trend of data.

The confidence interval calculated using t-distribution or Gaussian function for the data that obeys the normal distribution in short-term is estimated.

Figure 2: The illustration of time-domain filtering method [Shao (2016)]
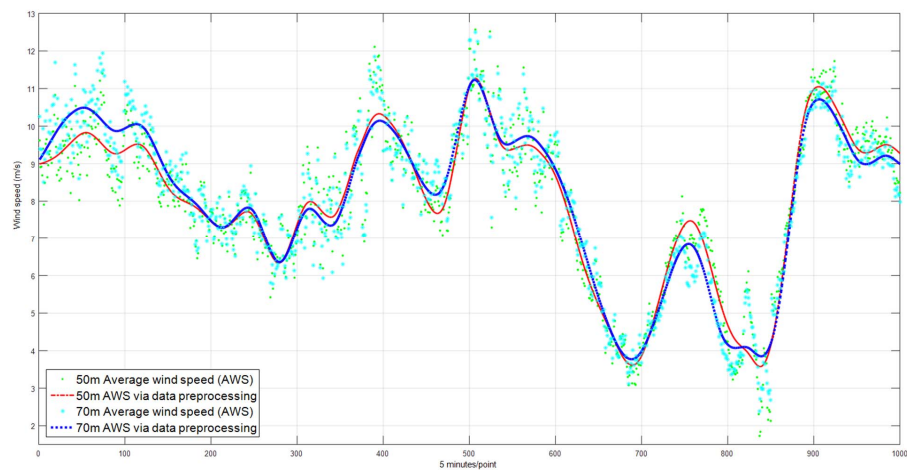
Figure 3: The illustration of time-domain filtering method [Shao, Wei, Deng et al. (2016)]

of variables is too small or the variables don't have sufficient information for forecasting. It can be intuitively judged whether the candidate variable has a reference value for the output variable prediction. In addition, the full use of a prior knowledge can not only effectively improve the computational efficiency, but also effectively promote the prediction accuracy.

(ii) Computational efficiency. Increase of the number of input variables corresponds to the increase of network volume and will decrease of the network learning efficiency. Take the kernel-based Generalization Regression Neural Network and Feedforward neural network as the examples. The metrics they used are highly susceptible to high-dimensional variables, which makes them computationally expensive.

(iii) Network learning efficiency. The training process of the network will become difficult when the input of the neural network contains redundant or irrelevant variables which result in an increase in the number of local optimal solutions in the error function. Assume that the first term on the right side of (1) is the error between the actual value and the forecasted value, while the second term is the error between the actual value and the forecasted value containing the redundant variable. Obviously, the estimated error is actually greater than the real one. The models forecasting error is given as follows,

$$
\begin{aligned}
Error_{y(t)} \approx & \frac{1}{k_N} \sqrt{\sum_{i=1,\ldots,k_N<k} \left(y_i(t) - \hat{y}_i(t)\right)^2} \\
& + \frac{1}{k-k_N} \sqrt{\sum_{j=k_N+1,\ldots,k} \left(y_j(t) - \hat{y}_j(t)\right)^2}
\end{aligned}
\tag{1}
$$

where $y_i(t)$ and $\hat{y}_i(t)$ are respectively the actual variables and their respective forecast. The error is regarded as the gradient direction of the learning process in neural network, since the influence on the learning process of the heuristic algorithm is significant. Take the widely used backpropagation algorithm based on Gradient Descent in neural network as an example, the network is susceptible to data distribution and easily converges to local extremum. In addition, the convergence speed of the network will be slower due to factors such as model redundancy parameters and noise. Especially, useless information may be added to adjust network weights, as well as other forecasts. As a result, many iterative algorithms require a near-global optimum error as constrains to supervise the learning algorithm.

(iv) Sample Dimensionality. The Curse of Dimensionality is a problem often encountered in the wind power forecasting modeling. The dimension of the model increases linearly on the surface, and the amount of model's calculation increases exponentially. In order to precisely estimate the model parameters, it is necessary to increase the number of samples in an exponential manner. In real applications, the established mapping may not fully reflect the trend of the actual data due to limited samples. In the case of the Multilayer Perceptron, due to the rapid increase in the number of samples, the Curse of Dimensionality will cause the network connection weights increase rapidly, which in turn leads to an increase in computation time and reduces its practicability.

The collected wind power time series can be treated as original features in forecast model at different time slots assuming its related large dimensionality and high degree of redundancy can be properly processed. Appropriate feature selection and analysis can be of great benefit to high-precision, high-efficiency wind power forecasting modeling. Taking the
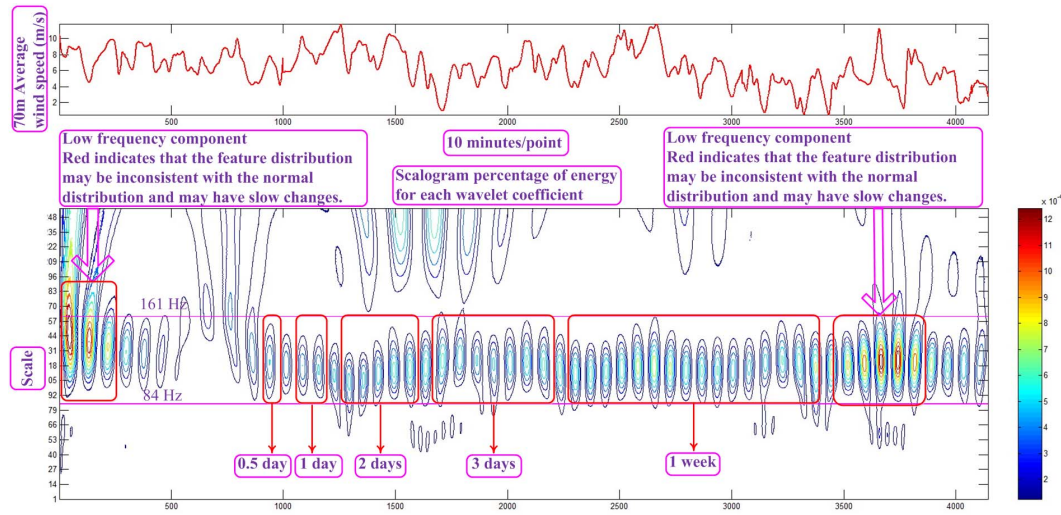
Figure 4: Seasonal pattern analysis [Shao, Wei, Deng et al. (2016)]

most widely used wavelet analysis as an example, the samples' wavelet coefficients in time-frequency domain can accurately reflect the its features' distribution, especially whether the training samples can provide effective information that used for the test samples to process the forecast modeling with high precision. Since the wind speed can be regarded as a cumulative superposition of different frequency components with volatility and periodicity, a wavelet transform based on short-time Fourier analysis as an effective approach is most widely used for analysis. Multi-layer decomposition of wind speed through multi-scale multi-resolution wavelet analysis can be used to find components with similar frequency characteristic resolution for wind speed, after which the forecasting model can be established to reflect the characteristics of the frequency components. Seasonal pattern analysis [Shao, Wei, Deng et al. (2016)] based on the wavelet analysis is shown in Fig. 4, which is mainly used to ensure the rationality of the division of the subsets of sample spaces.

3 Deep learning and reinforcement learning

The deep learning models with suitable architectures are designed to help select the inputs features [Dalto, Matuško and Vašak (2015); Chang, Lu, Chang et al. (2017)], and there are two types of models, deterministic and probabilistic forecasting models is established. However, due to lack of proper feature analysis, especially feature selection process which used for wind power forecasting, the aforementioned method usually requires a large amount of computation when processing high-dimensional, multi-input wind power data, which has a negative impact on scalability. Ensemble mixture density network (MDN) with three-layer architecture [Men, Yee, Lien et al. (2016)] is given for wind power forecasting, and the probabilistic forecasting measurement is applied to verify the experimental performance based on the data from the operational wind turbine in a wind

farm in Taiwan (72 h forecast horizon). The experimental evaluation demonstrates that the proposed methodology can achieve the better performance in multi-step ahead wind power forecasting. Different metric in learning algorithm are of great benefit to the acquisition of data features [Liu, Kwong and Chan (2012); Hossain, Rekabdar, Louis et al. (2015); Díaz, Torres and Dorronsoro (2015); Yan, Liu, Han et al. (2015); Luo, Sun, Wang et al. (2018)], for instance, the numerical simulation related to 1 h ahead forecasting at seven different locations of the surface radiation network indicates that the wind speed forecasting accuracy can be improved about 30% compared with the traditional benchmarks. For simplicity, the forecasting methods based on the deep learning and reinforcement learning etc is classified as the following categorized:

(i) RNN or LSTM-related approaches

RNN and LSTM accordingly predict the time sequence over time, which indicates that before the information entering the current processing unit, the long-term information of the input sequence needs to be pre-traversed by all hidden layer units in order. If the gradient associated with an hidden layer is minimal and prone to vanishing gradients, its improved version such as GRU (gated recurrent unit) [Fu, Hu, Tang et al. (2018)] is indeed beneficial to the solution of the aforementioned problems, in particular, for the longer-term sequence information, but its processing sequence is always limited (about 100 or so), in other words, the long-time information acquisition for longer sequences is still tricky. A novel deep learning approach based on the infinite feature selection (Inf-FS) with the recurrent neural networks (RNNs) is adopted in Shao et al. [Shao, Deng and Jiang (2018)]. The possible features related to wind forecast are first clustered into multiple feature sets, after which the identified feature sets are mapped onto the paths of a graph built for Inf-FS. Traversing such a graph helps effectively determine the significance of the features according to their stability and classification accuracy. Based on the data from the National Renewable Energy Laboratory (NREL), the experimental evaluation indicated that the short-term wind power forecasting accuracy is improved significantly by 11%, 29%, 33%, and 19% in spring, summer, autumn and winter, respectively. The one shortcoming of RNN is related to its vanishing gradient caused by the high power of the matrix. As a result, the long-term dependency of wind power time series cannot be established. Long Short-Term Memory (LSTM) network is one of the most widely used models in RNNs, which addresses the problems experienced in simple RNN [Wu and Lundstedt (1996); Senjyu, Yona, Urasaki et al. (2006); Barbounis, Theocharis, Alexiadis et al. (2006); Barbounis and Theocharis (2006, 2007); Aquino, Carvalho, Neto et al. (2010); Liu, Wu, Wang et al. (2018)], including the gradient exploding and gradient vanishing. In order to effectively analyze the mixture of long-term and short-term pattern of temporal series, LSTM, CNNs and RNNs are developed to extract the local dependency patterns of the inputs and discover the intrinsic pattern that used to describe the general trends [Lai, Chang, Yang et al. (2018)]. Experiments indicated that the proper analysis of the repetitive patterns of time series is beneficial to improve the final forecasting accuracy. LSTM network with a single hidden layer is shown in Fig. 5. LSTM network in combination with wavelet transformation as well as Elman neural network is adopted in Wu et al. [Wu, Chen, Qiao et al. (2016); López, Valle, Allende et al. (2017); Li, Zhu, Kong et al. (2018); Liu, Mi and Li (2018b); Dong,
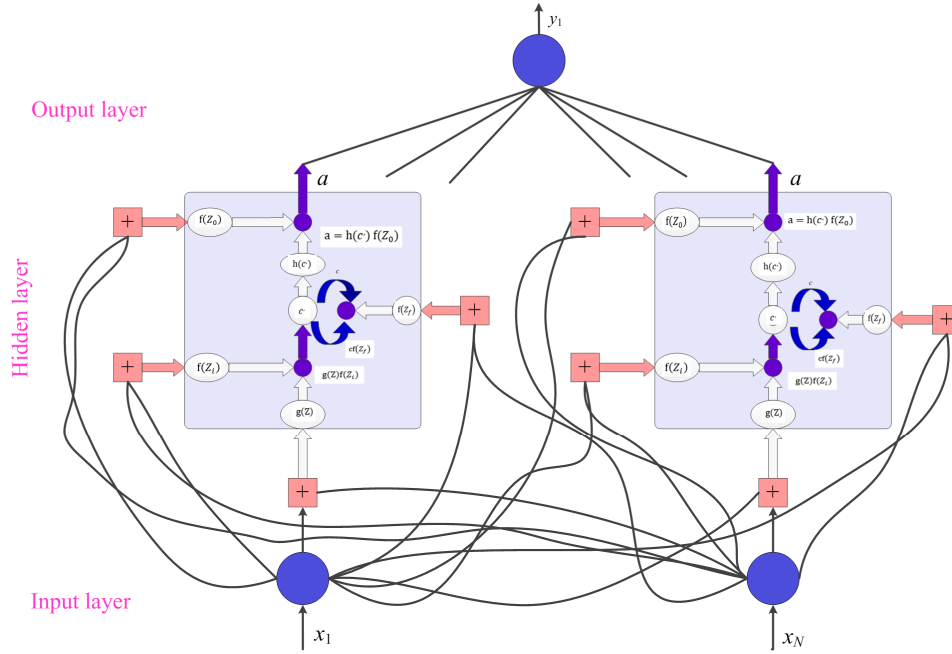
Figure 5: The LSTM network with single hidden layer

Sheng and Yang (2018); Liu, Mi and Li (2018a); Araya, Valle and Allende (2018); Guo and Lin (2018); Cao and Gui (2018); Liang, Nguyen and Jin (2018); Chen, Zhou, Li et al. (2018); Lei, Liu and Jiang (2019)] for multi-step wind speed forecasting, and the low-frequency and the high-frequency components obtained via the wavelet transformation are respectively forecasted by the LSTM network and Elman neural network with a hope to improve the total forecasting accuracy. The innovative short-term wind power forecasting system based on the ensemble methods and transfer learning is developed to construct the improved deep learning model [Qureshi, Khan, Zameer et al. (2017)].

(ii) ELM or SELM-related approaches

A novel bidirectional mechanism and a backward forecasting model based on extreme learning machine (ELM) are proposed to address the issue of ultra-short term wind power time series forecasting [Mohammadi, Shamshirband, Yee et al. (2015); Zhao, Ye, Li et al. (2016)]. Assuming that the previous measurement is unknown first, the reverse wind power time series are generated to train the backward ELM based on the optimization algorithm. The forecasting error between the forward results and backward ones are analyzed to adjust the model structure, optimize the model learning algorithm, and finally improve the forecasting accuracy. The experimental evaluation and comprehensive error analysis of 1-6 h ahead forecasting indicated that the bidirectional mechanism is beneficial to improve the wind power's forecasting accuracy. The stacked ELM (SELM) coupled with deep learning framework as well as less memory consumption are used to compare with the traditional ELM, which combined with correntropy are usually used to measure the similarity of inputs

and process the ultra-short-term wind speed forecasting. The computing performance indicated that the established model has better forecasting accuracy compared with other traditional methods. The uncertainty and instability of wind power time series always brings a challenge in high-precision forecasting accuracy [Tao, Chen and Qiu (2014); Paterakis, Mocanu, Gibescu et al. (2017)].

Different from the traditional feedforward neural network, ELM as a single hidden layer feedforward neural network only needs to set the number of hidden layers, and its input weight and hidden layer biases can be adjusted to obtain the unique optimal solution, with fast learning speed and good generalization ability. However, ELM is essentially a simple network based on the least norm least squares, which cannot utilize gradient-dependent chain rules used in DNNs. This indicates that it is not conducive to the extraction of high discriminative features of wind power data.

(iii) CNNs or DBNs-related approaches

Wavelet analysis as the most widely used time-frequency analysis method can be used to implement of the original time series' approximation with similar frequency characteristic. This is very beneficial to reduce the impact of non-stationarity and improve the accuracy of predictive modeling. As Sehnke proposed that the function fitting for the wind speed frequency distribution can be used to reduce the uncertainties of wind speed distribution [Sehnke, Strunk, Felder et al. (2013)], and this approaches can reflect the *state-of-the-art* in industrial aerodynamics related to wind resource assessment. The probabilistic wind power forecasting by using the deep learning ensemble approach is given in Wang et al. [Wang, Li, Wang et al. (2017)], and the time-frequency domain analysis and extraction of features with high discriminative for which are respectively processed by the wavelet transformation and convolution neural networks (CNNs) to achieve the purpose for which the data is accurately estimated. The competitive performance of the proposed approach shows that the uncertainties in wind power data can be accurately learned. Deep belief networks (DBNs) are utilized to forecast the wind power in Wang et al. [Wang, Wang, Li et al. (2016)], wavelet transformation, spine quantile regression (QR) and DBNs are separately used to analyze the non-stationary trend, statistically synthesized and extract the nonlinear features for the wind power data. The high-stability and exceptional performance of the proposed approaches are proved based on the experimental evaluation. Similarity, combined with variational model decomposition (VMD), the LSTM networks and ELM are respectively used to forecast the low-frequency and high-frequency component of the raw data. Four experiments denoted that the proposed hybrid approach can achieve the best forecasting accuracy compared with eight traditional models, and are beneficial to extract the effective and robust features as well as the trend information. Similar with the previous works, empirical model decomposition (EMD) [Hu, Wang and Zeng (2013)], wavelet transformations and multi-scale time-scale decomposition method are developed to reduce the negative influence from the non-stationary wind power time series in short-term. The data-driven multi-model for wind speed forecasting methodology is presented based on the ensemble neural networks.

(iv) Hybrid NNs-related approaches

A hybrid model related to short-term wind speed forecasting is presented based on the Wavelet Packet Decomposition, Convolutional Neural Network and Convolutional Long Short Term Memory Network (WPD-CNNLSTM-CNN) [Liu, Mi and Li (2018a)]. Designed to reduce the negative impact of non-stationary of raw data on short-term forecasting modeling, WPD is used to decompose the original wind speed time series into several sub-layers at differential levels. The CNN and CNNLSTM are respectively used to analyze the high-frequency and low-frequency of the time series and establish the forecasting model. Experimental evaluation shows that the hybrid modeling method for different frequencies can effectively improve the short-term wind speed forecasting accuracy on testing samples. Similarity, in order to overcome the non-stationary of wind power time series and analyze the wind speed's inherent volatility, the conditional mutual information, wavelet packet technique in combination with several individual models featured with different mixtures of mother wavelets are separately used to process the feature selection and construct the ensemble model [Li, Wang and Goel (2015); Meng, Ge, Yin et al. (2016)]. The multiple models are used to model the different frequency components of the data separately in order to fully extract the data characteristics [Ohlsson (2011); Liu, Tian, Liang et al. (2015); Mi, Liu and Li (2017); Wang, Zhang, Long et al. (2017); Liu, Mi and Li (2018a); Li, Xiao, Xia et al. (2018)], so as to improve the accuracy of short-term wind speed modeling. Wang et al. [Wang and Li (2018)] proposed an innovative hybrid approach for multi-step ahead wind speed forecasting based on the Kullback-Leibler divergence, energy measure, sample entropy in combination, to achieve the purpose of the optimal feature selection and improve the calculation efficiency. DNNs are finally used to capture the long-term and short-term memory data characteristics and establish the forecasting model. The forecasting error is analyzed by the generalized auto-regressive conditionally heteroscedastic model to update the evolutional information in modeling. Wind force model related to wind disturbances is developed to analyze the weather conditions, the simulation results of which indicates that the wind-induced vibrations and pointing errors of NASA Deep Space Network (DSN) antennas shares the high degree of consistency due to wind steady state pressure [Gawronski, Bienkiewicz and Hill (1994)]. Phase Space Reconstruction (PSR) and Kernel Principal Component Analysis (KPCA) are successively adopted to dynamically select the input vectors and effectively extract the nonlinear features of the original high-dimensional feature space reconstructed by PSR. Finally, Grey Relational Analysis as well as Pesaran-Timmermann statistic etc., is applied to assess the forecasting effectiveness of the proposed approaches. A comprehensive study related to wind speed forecasting interval from a few seconds to several months based on adaptive neuro-fuzzy inference system (ANFIS) and neural networks are given in Okumus et al. [Okumus and Dinler (2016); Chu, Yuan, Wang et al. (2019)]. The experiments indicate that absolute percentage errors (MAPE) based on the proposed approaches are 2.2598%, 3.3530% and 3.8589% at three different locations for average wind speeds daily. In order to make full use of the higher order correlations of the variants, DNNs is designed to meet the requirements of MCP for the wind resource assessment.

(v) Q-learning-related approaches

Deep learning has strong perceptual ability but with insufficient decision-making ability, and reinforcement learning has decision-making ability but lack of perceptual ability. Therefore, the effective combination of two aforementioned learning algorithms can achieve the goal of complementary advantages and enhance the generalization ability of the learning algorithm. Wei et al. [Wei, Zhang, Qiao et al. (2015)] proposed an intelligent maximum power point tracking training (MPPT) algorithm based on the reinforcement learning (RL). Q-learning method is applied based on the controller of the wind power conversion systems to establish the mapping between the system state and control actions online. The optimum wind-energy curve for maximum power points (MPPs) is then generated to control the wind energy conversion systems (WECSs) [Wei, Zhang, Qiao et al. (2015); Vandael, Claessens, Ernst et al. (2015); Wei, Zhang, Qiao et al. (2016); Wang, Zhang, Long et al. (2017)]. The model architecture and model parameters established based on the pre-learning process can effectively promote the interaction ability between wind energy conversion system and environment, so as to improve the accuracy of predictive modeling effectively. Salehizadeh et al. [Salehizadeh and Soltaniyan (2016)] proposed the fuzzy Q-learning approach for hour-ahead forecasting modeling related to the renewable resources. The experimental evaluation based on the IEEE 30-bus test system indicated that the proposed approach can model the inputs with continuous multi-dimensional variables and benefits the improvement of the computational efficiency. Zhang indicated that the combination-deep reinforcement learning (DRL) [Zhang, Han and Deng (2018)] is one of the most representative methods used to reduce the cost of the computing power and obtain the intrinsic patterns for the power systems, and has been widely used in smart grid and power system coordination control system. Xiao et al. [Xiao, Xiao, Dai et al. (2018)] proposed a microgrids (MGs) energy trading based on the reinforcement learning Q-network (DQN) to improve the utility of MGs, in particular, even in the case of a large number of connected MGs. Taking into account the MGs with wind generation, experiments indicated that the proposed scheme significantly reduce the energy consumption and benefit the improvement of the MG's utility compared with the traditional strategy. Hu et al. [Hu and Chen (2018)] proposed a hybrid scheme for wind speed forecasting based on the LSTM, ELM and evolution algorithm. The hysteresis added in the activation function of biological neural network to improve the generalization ability of ELM, and evolution algorithms are used to minimum the weighted objective function and derive the optimal number of the hidden layers of LSTM. Experiments based on four performance indices and statistical tests, denoted that compared with forecasting results derived by classical models, the ensemble outputs obtained by the proposed hybrid approach can show the superior advantages in predictive accuracy. Wind speed's stochastic and instantaneous characteristics hinder the extraction of its related variables in neighboring wind farms spatio-temporal features characteristics. A graph deep learning model is proposed by Khodayar et al. [Khodayar and Wang (2018)] to analyze the undirected graph related to wind site, which consisting of LSTM and graph convolutional deep learning architecture (GCDLA) are respectively to capture the inputs' temporal features and process the forecasting. The experimental results show that the hybrid deep learning methods based on different deep networks can effectively improve the prediction

accuracy of short-term wind speed. In addition, various improved versions of LSTM, for example [Xu and Xia (2018); Torres, Aguilar and Zuñiga-Meneses (2018); Lu, Sun, Wang et al. (2018); Yang, Wang, Aziz et al. (2018); Chen, Zeng, Lu et al. (2019); Zhang, Wang, Fang et al. (2018); Yu, Gao, Yu et al. (2019)], still have very high hardware requirements, and even if online learning process is not required, it still needs a lot of resources to quickly converge the network, that is, the improved version still does not match the hardware acceleration to form well. Transfer learning is beneficial to shorten the convergence time of the network and achieve the goal of quickly adapting to the new testing sample. The collective, robust and stable decision output by using the proposed approach on the testing data is derived to improve the generality ability of the deep learning. Because numerical weather prediction (NWP) fails to capture data's spatially characteristics sufficiently [Felder, Kaifel and Graves (2010); Colak, Sagiroglu and Yesilbudak (2012); Xiaoyun, Xiaoning, Chao et al. (2016); Wang, Qi, Liu et al. (2018); Dou, Liu and Wang (2018); Fu, Hu, Tang et al. (2018)], and the data-driven modeling always shows less adaptable on the new testing sample, the deep neural network in combination with the transfer learning framework [Hu, Zhang and Zhou (2016)] is given to extract a high-level representation of raw data and promote the model configuration in terms of rapid convergence. A hybrid architecture based on the revised bat algorithm (BA) with the conjugate gradient (CG) method is designed for multi-step ahead wind speed forecasting [Xiao, Qian and Shao (2017)]. Though the optimization of the weights initiations related to the hidden layer of deep neural networks (DNNs), the singular spectrum analysis in combination with general regression neural network achieved the better forecasting accuracy compared with the other existing methods. Because of the uncertainty and randomness of wind speed distribution, feature extraction with high discriminant based on the raw data is becoming more difficult, the predictive deep Boltzmann machine (PDBM) [Zhang, Chen, Gan et al. (2015)] is used to build the Long-term and short-term forecasting framework based on sophisticated deep-learning technique, then the probabilistic characteristics of wind speed are effectively captured based on the deep neural network.

4 Model structure optimization

Model structure optimization for deep learning is a very challenging problem. The polynomial-time algorithm does fit the convex optimization in simple neural network because all local extrema is usually treated as the global minimum [Zadeh and Goel (2013)], however, which is not suitable for deeper multi-layer neural networks. The completely reliable deep network optimization algorithm is almost non-existent because the model optimization problem is essentially Non-deterministic Polynomial hard also known as NP problem. In other words, it is not certain whether the currently trained deep network is the best model. High-precision learning is basically harmless when the training samples are free of noise. Since wind power data in wind farm even after data preprocessing still cannot completely determine whether all noise has been removed, excessively high learning accuracy forces the neural network to fit the noise contained in samples, resulting over-fitting in forecasting modeling. In addition, Due to the influence of the initial weight, the neural network has a certain randomness to achieve a given learning accuracy. In

practice, the most commonly used gradient descent optimization method is beneficial for obtaining sufficiently good local extrema, and heuristic method, regularization method, feature selection method, new hardware (such as GPU) and the use of gradient descent iterations as much as possible under limited time is of great benefit to improving the generalization ability of deep learning. In addition, the statistical performance measures based on the error analysis criteria, such as root mean squared error (RMSE) and mean absolute error (MAE) are widely used to compare with traditional existing techniques to verify the effectiveness of the proposed approaches in wind power forecasting. DBN is used to establish the nonlinear relationship between the historical data and variables needed to be forecasted. Deep learning with its powerful ability to capture the different patterns of wind power time series shows the significant value in both scientific and engineering application. In order to overcome the persistence statistics models disadvantages that caused by the randomness and uncontrollability of wind power times series, the deep neural network with three hidden layers in combination with stacked auto-encoders (SAE) is proposed in Jiao et al. [Jiao, Huang, Ma et al. (2018)]. The forecasting accuracy is better achieved based on the designed network architecture optimized by particle swarm optimization (PSO). The multi-step ahead forecasting results denoted that the forecasting accuracy of short-term wind power can be improved by 12%. A novel dynamical integrated approach is employed to implement the forecasting of wind speed and evaluate the potential assessment of wind energy [Sun, Qiao, Wei et al. (2017)]. The EnsemLSTM method [Chen, Zeng, Zhou et al. (2018)] is proposed by using LSTMs, Support Vector Regression (SVR) and Extremal Optimization (EO) Algorithm, to establish a nonlinear relationship between predictive output and historical data though the nonlinear-learning ensemble of DNNs. The cluster algorithm is developed to deal with various problems caused by data diversity, such as the slow transmission of information in deep network hidden layer and the inaccuracy of information expression. Experiments based on the data with sampling frequency 10 minutes from Inner Mongolia and China denoted that proposed EnsemLSTM can overcome the disadvantages of unstable prediction accuracy and low generalization ability of single model, and promote the generation of the forecasting model with high accuracy. Based on the outputs consisting of wind speed and daily solar radiation derived from NWP [Díaz-Vico, Torres-Barrán, Omari et al. (2017)], DNNs are used to capture and explore the intrinsic feature of the multi-dimensional inputs and performs of the state-of-the-art forecasting results.

5 Experiments

Architecture design, combination logic, application scenarios, such as coastal wind farms, offshore wind farms and inland wind farms, time scales as well as sampling frequencies etc are usually different in forecasting, this paper only compares some of the prediction method structures in the last three years. Based on the data derived from the sampling device weak wind turbine, named as type-FD77) in a wind farm plant of East China, Shao et al. [Shao, Deng and Cui (2016)] proposed a forecasting architecture based on AdaBoost neural networks in combination with wavelet decomposition to forecast the wind speed in short-term. The sampling frequency is 5 minutes/point, and 3 variables are selected

from the 18 variables such as average wind speed, average wind direction, real-time wind speed and real-time wind direction at 10 m, 50 m and 70 m, respectively, and 10 m air temperature, 10 m relative humidity and air pressure, and 24-steps (i.e., 2 h) ahead prediction is mainly considered. The average forecasting results related to TRA (traditional approach without model structure selection), FNN (Fuzzy model neural network), MSS (Model structure selection), MSS-Wav (MSS in combination with Wavelet decomposition at level Db4(2)), MSS-Ada (Neural network ensemble in combination with MSS-Wav) and HybridNN (Hybrid neural network) are given in Tab. 1.

Table 1: Results comparison

Seasons	Error	TRA	FNN	MSS	MSS-Wav	MSS-Ada	HybridNN
Spring	RMSE	1.2459	1.1579	1.1050	1.5819	1.1842	1.0543
	MAE	1.0153	0.9446	0.6217	0.5291	0.9042	0.8389
	RMAE	0.2293	0.2134	0.1404	0.1254	0.1176	0.1895
	ET	712.75	198.24	123.31	169.12	185.53	162.41
Summer	RMSE	1.6642	1.7174	1.8238	1.1632	1.1329	1.4357
	MAE	1.3030	1.3334	1.2292	0.7753	0.7520	1.0516
	RMAE	0.2306	0.2234	0.2176	0.1372	0.1331	0.1861
	ET	698.48	193.44	119.21	183.71	193.59	156.64
Autumn	RMSE	1.7374	1.7614	1.5308	2.9725	1.4658	1.9457
	MAE	1.3621	1.3209	1.1610	1.4638	1.0266	1.6032
	RMAE	0.2411	0.1979	0.2241	0.2193	0.1538	0.1904
	ET	736.18	172.52	142.21	176.14	156.68	138.47
Winter	RMSE	1.8238	1.2984	1.2141	1.1053	1.1096	1.5924
	MAE	1.2292	1.0159	0.9685	0.8886	0.6855	0.7090
	RMAE	0.2176	0.1878	0.1790	0.1642	0.1213	0.1311
	ET	729.19	169.59	136.89	189.78	149.65	128.56

where ET represents the elapsed time in seconds. The proposed approach is a classical network structure, i.e., multiple input and single output (MISO), and consisting of three layers, i.e., input layer, hidden layer and output layer. The number of the input layers is determined by the inputs selected by the model variables selection techniques etc. Of course, the considered system is actually not static, so the numbers of inputs layers changes over the time due to the seasonal nature of the wind speed. The average forecasting errors in terms of the results that reported in Tab. 1 indicated that the proposed approach MSS-Ada is reduces about TRA: 11.17%, 9.75%, 8.73% and 9.63%, FNN: 9.58%, 9.03%, 4.41% and 6.65%, MSS: 2.28%, 8.45%, 7.03% and 5.77%, MSS-Wav: 0.78%, 0.41%, 6.55% and 4.29%, HybridNN: 7.19%, 5.30%, 3.66% and 0.98%. The experiments denoted that the generalization ability of the model will be further improved if the irrelevant variables are removed, the model order is appropriately estimated and the wind speed frequency characteristics are properly considered in the final modeling. In Shao's later work [Shao, Wei, Deng et al. (2016)], the seasonal characteristics of the wind speed is more strictly considered and reflected through the dataset division. More precisely, the division of sample subsets is no longer the traditional equalization, but according to the seasonal

characteristics of wind speed distribution. This is more conducive to the training of samples and test samples of the characteristics analysis, and conducive to the mutual coverage of the spectrum. The 60%, 20% and 20% of the each divided subsets based on the data from the Yunnan wind farm are chosen to be the training sample, verification sample and testing sample, respectively. The experimetanal results related to the 12-steps (i.e., 2 h) ahead prediction is given in Tab. 2.

In Tab. 2, ET, TRA, TRD and MSN are respectively the computational time in seconds, traditional approach, TRA that considered the seasonal pattern and MSS without the neural network ensemble method. The sampling frequency is 10 minutes/point. Traditional approach is a traditional model with Single Input Single Output (SISO), in other words, the random and seasonal characteristics nature of the wind speed distribution cannot be considered in the forecaing. In fact, the AdaBoosting neural network combined with multi-individual models' learning abilities can siginficantly promote the network architecture synthesis, and then improve the generalization ability of the forecasting model. This approach can be treated as a forecasting model with good robustness and high accuracy, which has a range of engineering applicatons because it can reduce the errors caused by the artifically settting parameters etc. In order to effectively reflect the seasonal characteristics of the wind speed in forecasting modeling, there are $k = (6, 12, 18, 24)$ steps (corresponding to 1, 2, 3 and 4 h) ahead of actual wind power forecasting is also proposed, and the averaged forecasting results obtained for each season based on the data division derived by the seasonal pattern are tabulated in Tab. 3. FS: specified steps ahead forecasting; RMSE1, RSD1, RMSE2 and RSD2 are the RMSE (Root mean square error) and RSD (Relative standard deviation) of the LM (Levenberg-Marquardt) output and BFGS (Quasi-Newton methods) output in the training. RMSE3, RSD3, RMSE4 and RSD4 are the RMSE and RSD of the LM output and BFGS output in the testing. ET indicates the elapsed time in seconds. Although the high predication accuracy of the proposed approach comes at a high price of elapsed time, it is still a robustness approach for the wind power forecasting. Usually, the forecasting performance deteriorates along with the increment of the forecasting-steps. In particular, the small error in wind speed forecasting usually causes the big errors in wind power forecasting. The forecasting model should have the ability of the error correction, dynamical feedback and adaptive adjustment in real application. The forecasting results obtained by the proposed method are compared with other two classical methods, i.e., persistence and new reference, which indicated that the forecasting errors MAPE related to persistence model about the 1 hour-ahead, 24 h-ahead and 48 h-ahead prediction is respectively reduced by 93.74%, 82.04% and 82.46%. In the case of the new reference, MAPE is respectively reduced by 94.10%, 77.51% and 77.92%. As we mentioned before, with the powerful capability of wavelet transformation in time-frequency domain analysis, the instantaneous, randomness and non-stationary of wind speed distribution in short-term are greatly reduced, which is of great benefit to the improvement of prediction accuracy. As the Zhang et al. [Zhang, Chen, Gan et al. (2015)] reported, the 1-h ahead forecasting results of five feature selection methods is given in Tab. 5, Taking into account the CMIFS, the forecasting accuracy is improved by 15.5%, 21.0%, 19.7% and 67.5% in MAPE refer to the preceding four methods, respectively.

Table 2: Forecasting results

Seasons	Error	TRA	DS	TRD	MSN	PRA
Spring	RMSE	1.6084	8	1.5911	1.0224	1.0448
	MAE	1.2901		1.2324	0.7501	0.7722
	RMAE	0.2251		0.1316	0.0801	0.0790
	ET	43.53		30.87	28.04	42.91
	RMSE		9	1.5366	0.9691	1.0269
	MAE			1.1553	0.7277	0.7058
	RMAE			0.2122	0.1462	0.1418
	ET			32.57	33.09	41.49
Summer	RMSE	1.4722	1	1.3859	1.3884	1.3216
	MAE	1.1792		1.1766	1.0506	0.9943
	RMAE	0.2596		0.2137	0.2091	0.1978
	ET	43.37		32.62	34.15	42.72
	RMSE		2	1.3277	0.7317	0.7447
	MAE			1.0845	0.5349	0.5626
	RMAE			0.2434	0.1200	0.1162
	ET			39.44	38.32	42.45
Autumn	RMSE	1.3804	3	1.0938	0.8319	0.7258
	MAE	1.0340		0.8983	0.5804	0.5250
	RMAE	0.1713		0.1416	0.0915	0.0827
	ET	43.72		39.52	35.40	43.36
	RMSE		4	1.7275	1.0163	1.1172
	MAE			1.3032	0.7540	0.7625
	RMAE			0.1916	0.1109	0.1102
	ET			32.51	34.29	42.08
Winter	RMSE	1.5899	5	1.5729	1.0567	1.0547
	MAE	1.2604		1.2074	0.8128	0.7849
	RMAE	0.1168		0.1098	0.0897	0.0866
	ET	43.46		32.34	32.90	41.21
	RMSE		6	1.4374	1.0141	0.8960
	MAE			1.1040	0.7709	0.6712
	RMAE			0.0964	0.0673	0.0586
	ET			32.39	32.84	39.24
	RMSE		7	1.5479	1.2444	1.0838
	MAE			1.1865	0.8597	0.7235
	RMAE			0.1011	0.0732	0.0616
	ET			32.47	32.84	37.84

Table 3: Forecasting results based on the proposed approach

	FS	ET	RM1	RS1	RM2	RS2	RM3	RS3	RM4	RS4
Spring	6	2618	0.16	1.88	0.18	2.08	0.17	2.61	0.17	2.61
	12	2628	0.23	2.59	0.24	2.73	0.54	4.37	0.48	4.18
	18	2590	0.29	3.24	0.35	3.96	0.48	5.15	0.48	5.15
	24	2620	0.32	3.57	0.34	3.84	0.49	5.18	0.49	5.18
Summer	6	2562	0.11	1.34	0.12	1.50	0.11	1.40	0.11	1.40
	12	2559	0.24	2.83	0.26	3.14	0.30	3.78	0.30	3.78
	18	2568	0.29	3.43	0.32	3.70	0.44	5.44	0.44	5.44
	24	2619	0.34	3.98	0.36	4.30	0.43	5.38	0.43	5.38
Autumn	6	2577	0.17	1.63	0.19	1.90	0.15	2.03	0.15	2.03
	12	2564	0.34	3.32	0.42	3.86	0.35	3.48	0.35	3.48
	18	2588	0.39	3.75	0.42	4.08	0.53	5.26	0.53	5.26
	24	2661	0.46	4.46	0.47	4.54	0.63	7.65	0.63	7.65
Winter	6	2725	0.18	2.04	0.23	2.49	0.17	3.86	0.17	3.86
	12	2668	0.28	3.04	0.30	3.29	0.43	5.54	0.43	5.54
	18	2629	0.33	3.60	0.33	3.64	0.58	6.49	0.58	6.49
	24	2613	0.34	3.77	0.38	4.21	0.57	7.37	0.57	7.37

Table 4: Forecasting results comparision [Zhang, Chen, Gan et al. (2015)]

Methods	Criteria	1hour	24hours	48hours
Persistence	MAPE	7.83	26.90	30.67
	NRMSE	12.5	35.50	39.21
New reference	MAPE	8.30	21.48	24.37
	NRMSE	12.26	23.86	26.93
Proposed	MAPE	0.49	4.83	5.38
	NRMSE	0.66	6.29	6.98

Table 5: 1-h ahead forecasting results of five feature selection methods

Methods	MAPE	NRMSE
CA	0.58	0.83
MI	0.62	0.91
RReliefF	0.61	0.87
mRMR	1.51	0.18
CMIFS	0.49	0.66

In fact, the forecasting accuracy is impossible 100% correct, forecasting model structure design methods need to take full account of the distribution characteristics of wind power samples, seasonal characteristics and dynamic adjustment of model capacity, especially for deep learning and online learning, in order to better improve the forecasting accuracy of prediction and adapt to practical engineering applications. Similar with the forecasting

Table 6: Deterministic 1-h ahead forecasting error at SIWF

Season	Error	ARMA	BPNN	MWNN	Proposed
Spring	MAE	0.6183	0.6330	0.6869	0.3609
	RMSE	0.8009	0.8152	0.8549	0.4771
	MAPE	11.02%	11.28%	12.24%	6.43%
Summer	MAE	0.6719	0.7850	0.7392	0.3415
	RMSE	0.9328	1.0685	0.9647	0.4391
	MAPE	10.94%	12.78%	12.03%	5.56%
Fall	MAE	1.0606	1.2232	1.1396	0.4947
	RMSE	1.3935	1.5917	1.5143	0.6671
	MAPE	10.20%	11.77%	10.96%	4.76%
Winter	MAE	0.7176	0.7166	0.7583	0.5155
	RMSE	0.9304	0.9309	0.9722	0.6142
	MAPE	12.27%	12.25%	12.96%	8.81%
Average	MAE	0.7671	0.8394	0.8310	0.4282
	RMSE	1.0144	1.1016	1.0765	0.5494
	MAPE	11.11%	12.02%	12.05%	6.39%

architecture used in Zhang et al. [Zhang, Chen, Gan et al. (2015)] work, the forecasting model combined with the wavelet transformation and deep learning approach is still used in wind speed forecasting by Wang et al. [Wang, Wang, Li et al. (2016)]. According to the deterministic 1-h ahead forecasting error, the forecasting error in four seasons and average results is reduced by 6.43%, 5.56%, 4.76%, 8.81% and 6.39%. This also indicated the hybrid approach which can effectively combine the advantages of different methods can effectively improve the prediction accuracy of wind power.

6 Conclusions

In this paper, the wind power forecasting modeling based on deep learning was formulated and discussed. The fundamental forecasting frameworks related to the deep learning-related methods, consisting of multi-feature and single-feature selection are firstly given. The literature comparison of the deep learning-related approaches, and the advantages and disadvantages of various related methods as well as the improvement strategy are then discussed in detail. The performance evaluation criteria and model optimization strategy for deep learning-related approaches are finally provided. We believe that the proposed strategy in this paper is beneficial to improve the relevant researchers and engineers' understanding and application in architecture design, model structure selection

and optimization in wind power forecasting modeling based on deep learning-related approaches.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

An, X.; Jiang, D.; Li, S.; Zhao, M. (2011): Application of the ensemble empirical mode decomposition and hilbert transform to pedestal looseness study of direct-drive wind turbine. *Energy*, vol. 36, no. 9, pp. 5508-5520.

Aquino, R. R.; Carvalho, M. A.; Neto, O. N.; Lira, M. M.; de Almeida, G. J. et al. (2010): Recurrent neural networks solving a real large scale mid-term scheduling for power plants. *International Joint Conference on Neural Networks*, pp. 1-6.

Araya, I. A.; Valle, C.; Allende, H. (2018): LSTM-based multi-scale model for wind speed forecasting. *Iberoamerican Congress on Pattern Recognition*, pp. 38-45.

Barbounis, T.; Theocharis, J. (2006): Locally recurrent neural networks for long-term wind speed and power prediction. *Neurocomputing*, vol. 69, no. 4-6, pp. 466-496.

Barbounis, T.; Theocharis, J. B. (2007): Locally recurrent neural networks for wind speed prediction using spatial correlation. *Information Sciences*, vol. 177, no. 24, pp. 5775-5797.

Barbounis, T. G.; Theocharis, J. B.; Alexiadis, M. C.; Dokopoulos, P. S. (2006): Long-term wind speed and power forecasting using local recurrent neural network models. *IEEE Transactions on Energy Conversion*, vol. 21, no. 1, pp. 273-284.

Cao, Y.; Gui, L. (2018): Multi-step wind power forecasting model using LSTM networks, similar time series and lightGBM. *5th International Conference on Systems and Informatics*, pp. 192-197.

Chang, G.; Lu, H.; Chang, Y.; Lee, Y. (2017): An improved neural network-based approach for short-term wind speed and power forecast. *Renewable Energy*, vol. 105, pp. 301-311.

Chen, C.; Zhou, J.; Li, X.; Zhao, P.; Peng, T. et al. (2018): Ultrashort-term forecasting of mid-low-altitude wind based on LSTM. *China High Resolution Earth Observation Conference*, pp. 104-111.

Chen, J.; Zeng, G. Q.; Zhou, W.; Du, W.; Lu, K. D. (2018): Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization. *Energy Conversion and Management*, vol. 165, pp. 681-695.

Chen, M. R.; Zeng, G. Q.; Lu, K. D.; Weng, J. (2019): A two-layer nonlinear combination method for short-term wind speed prediction based on ELM, ENN and LSTM. *IEEE Internet of Things Journal*.

Chu, J.; Yuan, L.; Wang, W.; Pan, L.; Wei, J. (2019): Short-term wind speed prediction based on deep learning. *IOP Conference Series: Earth and Environmental Science*, vol. 233, no. 5, 052007.

Colak, I.; Sagiroglu, S.; Yesilbudak, M. (2012): Data mining and wind power prediction: a literature review. *Renewable Energy*, vol. 46, pp. 241-247.

Dalto, M.; Matuško, J.; Vašak, M. (2015): Deep neural networks for ultra-short-term wind forecasting. *IEEE International Conference on Industrial Technology*, pp. 1657-1663.

Díaz, D.; Torres, A.; Dorronsoro, J. R. (2015): Deep neural networks for wind energy prediction. *International Work-Conference on Artificial Neural Networks*, pp. 430-443.

Díaz-Vico, D.; Torres-Barrán, A.; Omari, A.; Dorronsoro, J. R. (2017): Deep neural networks for wind and solar energy prediction. *Neural Processing Letters*, vol. 46, no. 3, pp. 829-844.

Dong, D.; Sheng, Z.; Yang, T. (2018): Wind power prediction based on recurrent neural network with long short-term memory units. *International Conference on Renewable Energy and Power Engineering*, pp. 34-38.

Dou, J.; Liu, C.; Wang, B. (2018): Short-term wind power forecasting based on convolutional neural networks. *IOP Conference Series: Earth and Environmental Science*, vol. 170, no. 4, 042023.

Felder, M.; Kaifel, A.; Graves, A. (2010): Wind power prediction using mixture density recurrent neural networks. *Poster Presentation Gehalten AUF DER European Wind Energy Conference*.

Feng, Z.; Liang, M.; Zhang, Y.; Hou, S. (2012): Fault diagnosis for wind turbine planetary gearboxes via demodulation analysis based on ensemble empirical mode decomposition and energy separation. *Renewable Energy*, vol. 47, pp. 112-126.

Fu, Y.; Hu, W.; Tang, M.; Yu, R.; Liu, B. (2018): Multi-step ahead wind power forecasting based on recurrent neural networks. *IEEE PES Asia-Pacific Power and Energy Engineering Conference*, pp. 217-222.

Gawronski, W.; Bienkiewicz, B.; Hill, R. (1994): Wind-induced dynamics of a deep space network antenna. *Journal of Sound and Vibration*, vol. 178, no. 1, pp. 67-77.

Guo, T.; Lin, T. (2018): Multi-variable lstm neural network for autoregressive exogenous model. arxiv:1806.06384.

Hodge, B. M.; Milligan, M. (2011): Wind power forecasting error distributions over multiple timescales. *IEEE Power and Energy Society General Meeting*, pp. 1-8.

Hong, W. C.; Li, M. W.; Geng, J.; Zhang, Y. (2019): Novel chaotic bat algorithm for forecasting complex motion of floating platforms. *Applied Mathematical Modelling*, vol. 72, pp. 425-443.

Hossain, M.; Rekabdar, B.; Louis, S. J.; Dascalu, S. (2015): Forecasting the weather of nevada: a deep learning approach. *International Joint Conference on Neural Networks*, pp. 1-6.

Hu, J.; Wang, J.; Zeng, G. (2013): A hybrid forecasting approach applied to wind speed time series. *Renewable Energy*, vol. 60, pp. 185-194.

Hu, Q.; Zhang, R.; Zhou, Y. (2016): Transfer learning for short-term wind speed prediction with deep neural networks. *Renewable Energy*, vol. 85, pp. 83-95.

Hu, Y. L.; Chen, L. (2018): A nonlinear hybrid wind speed forecasting model using lstm network, hysteretic elm and differential evolution algorithm. *Energy Conversion and Management*, vol. 173, pp. 123-142.

Jiao, R.; Huang, X.; Ma, X.; Han, L.; Tian, W. (2018): A model combining stacked auto encoder and back propagation algorithm for short-term wind power forecasting. *IEEE Access*, vol. 6, pp. 17851-17858.

Khodayar, M.; Wang, J. (2018): Spatio-temporal graph deep neural network for short-term wind speed forecasting. *IEEE Transactions on Sustainable Energy*, vol. 10, no. 2, pp. 670-681.

Lahmiri, S. (2015): Comparing variational and empirical mode decomposition in forecasting day-ahead energy prices. *IEEE Systems Journal*, vol. 11, no. 3, pp. 1907-1910.

Lai, G.; Chang, W. C.; Yang, Y.; Liu, H. (2018): Modeling long-and short-term temporal patterns with deep neural networks. *41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, pp. 95-104.

Lei, J.; Liu, C.; Jiang, D. (2019): Fault diagnosis of wind turbine based on long short-term memory networks. *Renewable Energy*, vol. 133, pp. 422-432.

Li, C.; Xiao, Z.; Xia, X.; Zou, W.; Zhang, C. (2018): A hybrid model based on synchronous optimisation for multi-step short-term wind speed forecasting. *Applied Energy*, vol. 215, pp. 131-144.

Li, S.; Wang, P.; Goel, L. (2015): Wind power forecasting using neural network ensembles with feature selection. *IEEE Transactions on Sustainable Energy*, vol. 6, no. 4, pp. 1447-1456.

Li, Y.; Zhu, Z.; Kong, D.; Han, H.; Zhao, Y. (2018): EA-LSTM: evolutionary attention-based LSTM for time series prediction. arxiv:1811.03760.

Li, Z.; Jiang, Y.; Guo, Q.; Hu, C.; Peng, Z. (2018): Multi-dimensional variational mode decomposition for bearing-crack detection in wind turbines with large driving-speed variations. *Renewable Energy*, vol. 116, pp. 55-73.

Liang, S.; Nguyen, L.; Jin, F. (2018): A multi-variable stacked long-short term memory network for wind speed forecasting. *IEEE International Conference on Big Data*, pp. 4561-4564.

Liu, H.; Chen, C.; Tian, H. Q.; Li, Y. F. (2012): A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. *Renewable Energy*, vol. 48, pp. 545-556.

Liu, H.; Mi, X.; Li, Y. (2018): Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolutional neural network and convolutional long short term memory network. *Energy Conversion and Management*, vol. 166, pp. 120-131.

Liu, H.; Mi, X. W.; Li, Y. F. (2018): 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.

Liu, H.; Tian, H. Q.; Liang, X. F.; Li, Y. F. (2015): Wind speed forecasting approach using secondary decomposition algorithm and elman neural networks. *Applied Energy*, vol. 157, pp. 183-194.

Liu, J. N.; Kwong, K.; Chan, P. (2012): Chaotic oscillatory-based neural network for wind shear and turbulence forecast with lidar data. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, pp. 1412-1423.

Liu, T.; Wu, T.; Wang, M.; Fu, M.; Kang, J. et al. (2018): Recurrent neural networks based on LSTM for predicting geomagnetic field. *IEEE International Conference on Aerospace Electronics and Remote Sensing Technology*, pp. 1-5.

López, E.; Valle, C.; Allende, H.; Gil, E. (2017): Long short-term memory networks based in echo state networks for wind speed forecasting. *Iberoamerican Congress on Pattern Recognition*, pp. 347-355.

Lu, K.; Sun, W. X.; Wang, X.; Meng, X. R.; Zhai, Y. et al. (2018): Short-term wind power prediction model based on encoder-decoder LSTM. *IOP Conference Series: Earth and Environmental Science*, vol. 186, no. 5, 012020.

Luo, X.; Sun, J.; Wang, L.; Wang, W.; Zhao, W. et al. (2018): Short-term wind speed forecasting via stacked extreme learning machine with generalized correntropy. *IEEE Transactions on Industrial Informatics*, vol. 14, no. 11, pp. 4963-4971.

Men, Z.; Yee, E.; Lien, F. S.; Wen, D.; Chen, Y. (2016): Short-term wind speed and power forecasting using an ensemble of mixture density neural networks. *Renewable Energy*, vol. 87, pp. 203-211.

Meng, A.; Ge, J.; Yin, H.; Chen, S. (2016): Wind speed forecasting based on wavelet packet decomposition and artificial neural networks trained by crisscross optimization algorithm. *Energy Conversion and Management*, vol. 114, pp. 75-88.

Mi, X. W.; Liu, H.; Li, Y. F. (2017): Wind speed forecasting method using wavelet, extreme learning machine and outlier correction algorithm. *Energy Conversion and Management*, vol. 151, pp. 709-722.

Mohammadi, K.; Shamshirband, S.; Yee, L.; Petković, D.; Zamani, M. et al. (2015): Predicting the wind power density based upon extreme learning machine. *Energy*, vol. 86, pp. 232-239.

Naik, J.; Dash, S.; Dash, P.; Bisoi, R. (2018): Short term wind power forecasting using hybrid variational mode decomposition and multi-kernel regularized pseudo inverse neural network. *Renewable Energy*, vol. 118, pp. 180-212.

Ohlsson, S. (2011): *Deep Learning: How the Mind Overrides Experience*. Cambridge University Press.

Okumus, I.; Dinler, A. (2016): Current status of wind energy forecasting and a hybrid method for hourly predictions. *Energy Conversion and Management*, vol. 123, pp. 362371.

Pant, T.; Han, C.; Wang, H. (2019): Examination of errors of table integration in flamelet/progress variable modeling of a turbulent non-premixed jet flame. *Applied Mathematical Modelling*, vol. 72, pp. 369-384.

Paterakis, N. G.; Mocanu, E.; Gibescu, M.; Stappers, B.; van Alst, W. (2017): Deep learning versus traditional machine learning methods for aggregated energy demand prediction. *IEEE PES Innovative Smart Grid Technologies Conference Europe*, pp. 1-6.

Qureshi, A. S.; Khan, A.; Zameer, A.; Usman, A. (2017): Wind power prediction using deep neural network based meta regression and transfer learning. *Applied Soft Computing*, vol. 58, pp. 742-755.

Ren, Y.; Suganthan, P.; Srikanth, N. (2014): A comparative study of empirical mode decomposition-based short-term wind speed forecasting methods. *IEEE Transactions on Sustainable Energy*, vol. 6, no. 1, pp. 236-244.

Ren, Y.; Suganthan, P. N.; Srikanth, N. (2014): A novel empirical mode decomposition with support vector regression for wind speed forecasting. *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 8, pp. 1793-1798.

Salehizadeh, M. R.; Soltaniyan, S. (2016): Application of fuzzy q-learning for electricity market modeling by considering renewable power penetration. *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 1172-1181.

Sehnke, F.; Strunk, A.; Felder, M.; Brombach, J.; Kaifel, A. et al. (2013): Wind power resource estimation with deep neural networks. *International Conference on Artificial Neural Networks*, pp. 563-570.

Senjyu, T.; Yona, A.; Urasaki, N.; Funabashi, T. (2006): Application of recurrent neural network to long-term-ahead generating power forecasting for wind power generator. *IEEE PES Power Systems Conference and Exposition*, pp. 1260-1265.

Shao, H. (2016): *Short-term Wind Speed Forecasting Based on Data in Wind Farm.* Southeast University.

Shao, H.; Deng, X. (2018): Adaboosting neural network for short-term wind speed forecasting based on seasonal characteristics analysis and lag space estimation. *Computer Modeling in Engineering & Sciences*, vol. 114, no. 3, pp. 277-293.

Shao, H.; Deng, X.; Cui, F. (2016): Short-term wind speed forecasting using the wavelet decomposition and adaboost technique in wind farm of east china. *IET Generation, Transmission & Distribution*, vol. 10, no. 11, pp. 2585-2592.

Shao, H.; Deng, X.; Jiang, Y. (2018): A novel deep learning approach for short-term wind power forecasting based on infinite feature selection and recurrent neural network. *Journal of Renewable and Sustainable Energy*, vol. 10, no. 4, 043303.

Shao, H.; Wei, H.; Deng, X.; Xing, S. (2016): Short-term wind speed forecasting using wavelet transformation and adaboosting neural networks in yunnan wind farm. *IET Renewable Power Generation*, vol. 11, no. 4, pp. 374-381.

Sun, S.; Qiao, H.; Wei, Y.; Wang, S. (2017): A new dynamic integrated approach for wind speed forecasting. *Applied Energy*, vol. 197, pp. 151-162.

Sun, W.; Wang, Y. (2018): Short-term wind speed forecasting based on fast ensemble empirical mode decomposition, phase space reconstruction, sample entropy and improved back-propagation neural network. *Energy Conversion and Management*, vol. 157, pp. 1-12.

Tao, Y.; Chen, H.; Qiu, C. (2014): Wind power prediction and pattern feature based on deep learning method. *IEEE PES Asia-Pacific Power and Energy Engineering Conference*, pp. 1-4.

Torres, J.; Aguilar, R.; Zuñiga-Meneses, K. (2018): Deep learning to predict the generation of a wind farm. *Journal of Renewable and Sustainable Energy*, vol. 10, no. 1, 013305.

Vandael, S.; Claessens, B.; Ernst, D.; Holvoet, T.; Deconinck, G. (2015): Reinforcement learning of heuristic ev fleet charging in a day-ahead electricity market. *IEEE Transactions on Smart Grid*, vol. 6, no. 4, pp. 1795-1805.

Wang, D.; Luo, H.; Grunder, O.; Lin, Y. (2017): Multi-step ahead wind speed forecasting using an improved wavelet neural network combining variational mode decomposition and phase space reconstruction. *Renewable Energy*, vol. 113, pp. 1345-1358.

Wang, H.; Wang, G.; Li, G.; Peng, J.; Liu, Y. (2016): Deep belief network based deterministic and probabilistic wind speed forecasting approach. *Applied Energy*, vol. 182, pp. 80-93.

Wang, H. Z.; Li, G. Q.; Wang, G. B.; Peng, J. C.; Jiang, H. et al. (2017): Deep learning based ensemble approach for probabilistic wind power forecasting. *Applied Energy*, vol. 188, pp. 56-70.

Wang, J.; Li, Y. (2018): Multi-step ahead wind speed prediction based on optimal feature extraction, long short term memory neural network and error correction strategy. *Applied Energy*, vol. 230, pp. 429-443.

Wang, J.; Zhang, W.; Li, Y.; Wang, J.; Dang, Z. (2014): Forecasting wind speed using empirical mode decomposition and elman neural network. *Applied Soft Computing*, vol. 23, pp. 452-459.

Wang, K.; Ji, C.; Xue, H.; Tang, W. (2016): Fatigue damage characteristics of a semisubmersible-type floating offshore wind turbine at tower base. *Journal of Renewable and Sustainable Energy*, vol. 8, no. 5, 053307.

Wang, K.; Qi, X.; Liu, H.; Song, J. (2018): Deep belief network based k-means cluster approach for short-term wind power forecasting. *Energy*, vol. 165, pp. 840-852.

Wang, L.; Zhang, Z.; Long, H.; Xu, J.; Liu, R. (2017): Wind turbine gearbox failure identification with deep neural networks. *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 1360-1368.

Wang, L.; Zhang, Z.; Xu, J.; Liu, R. (2018): Wind turbine blade breakage monitoring with deep autoencoders. *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 2824-2833.

Wang, Q.; Liu, W.; Yu, H.; Zheng, S.; Gao, S. et al. (2019): CPAC: energy-efficient algorithm for iot sensor networks based on enhanced hybrid intelligent swarm. *Computer Modeling in Engineering & Sciences*, vol. 121, no. 1, pp. 83-103.

Wang, S.; Zhang, N.; Wu, L.; Wang, Y. (2016): Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and ga-bp neural network method. *Renewable Energy*, vol. 94, pp. 629-636.

Wei, C.; Zhang, Z.; Qiao, W.; Qu, L. (2015): Reinforcement-learning-based intelligent maximum power point tracking control for wind energy conversion systems. *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10, pp. 6360-6370.

Wei, C.; Zhang, Z.; Qiao, W.; Qu, L. (2016): An adaptive network-based reinforcement learning method for MPPT control of PMSG wind energy conversion systems. *IEEE Transactions on Power Electronics*, vol. 31, no. 11, pp. 7837-7848.

Wu, J. G.; Lundstedt, H. (1996): Prediction of geomagnetic storms from solar wind data using elman recurrent neural networks. *Geophysical Research Letters*, vol. 23, no. 4, pp. 319-322.

Wu, W.; Chen, K.; Qiao, Y.; Lu, Z. (2016): Probabilistic short-term wind power forecasting based on deep neural networks. *International Conference on Probabilistic Methods Applied to Power Systems*, pp. 1-8.

Wu, X.; Zhu, Z.; Su, X.; Fan, S.; Du, Z. et al. (2015): A study of single multiplicative neuron model with nonlinear filters for hourly wind speed prediction. *Energy*, vol. 88, pp. 194-201.

Xiao, L.; Qian, F.; Shao, W. (2017): Multi-step wind speed forecasting based on a hybrid forecasting architecture and an improved bat algorithm. *Energy Conversion and Management*, vol. 143, pp. 410-430.

Xiao, L.; Xiao, X.; Dai, C.; Pengy, M.; Wang, L. et al. (2018): Reinforcement learning-based energy trading for microgrids. arxiv:1801.06285.

Xiaoyun, Q.; Xiaoning, K.; Chao, Z.; Shuai, J.; Xiuda, M. (2016): Short-term prediction of wind power based on deep long short-term memory. *IEEE PES Asia-Pacific Power and Energy Engineering Conference*, pp. 1148-1152.

Xu, G.; Xia, L. (2018): Short-term prediction of wind power based on adaptive LSTM. *2nd IEEE Conference on Energy Internet and Energy System Integration (EI2)*, pp. 1-5.

Yan, J.; Liu, Y.; Han, S.; Wang, Y.; Feng, S. (2015): Reviews on uncertainty analysis of wind power forecasting. *Renewable and Sustainable Energy Reviews*, vol. 52, pp. 1322-1330.

Yang, T.; Wang, H.; Aziz, S.; Jiang, H.; Peng, J. (2018): A novel method of wind speed prediction by peephole LSTM. *International Conference on Power System Technology*, pp. 364-369.

Yang, W.; Peng, Z.; Wei, K.; Shi, P.; Tian, W. (2016): Superiorities of variational mode decomposition over empirical mode decomposition particularly in time-frequency feature extraction and wind turbine condition monitoring. *IET Renewable Power Generation*, vol. 11, no. 4, pp. 443-452.

Yeh, J. R.; Shieh, J. S.; Huang, N. E. (2010): Complementary ensemble empirical mode decomposition: a novel noise enhanced data analysis method. *Advances in Adaptive Data Analysis*, vol. 2, no. 2, pp. 135-156.

Yu, R.; Gao, J.; Yu, M.; Lu, W.; Xu, T. et al. (2019): LSTM-EFG for wind power forecasting based on sequential correlation features. *Future Generation Computer Systems*, vol. 93, pp. 33-42.

Zadeh, R. B.; Goel, A. (2013): Dimension independent similarity computation. *Journal of Machine Learning Research*, vol. 14, no. 1, pp. 1605-1626.

Zhang, C. Y.; Chen, C. P.; Gan, M.; Chen, L. (2015): Predictive deep boltzmann machine for multiperiod wind speed forecasting. *IEEE Transactions on Sustainable Energy*, vol. 6, no. 4, pp. 1416-1425.

Zhang, D.; Han, X.; Deng, C. (2018): Review on the research and practice of deep learning and reinforcement learning in smart grids. *CSEE Journal of Power and Energy Systems*, vol. 4, no. 3, pp. 362-370.

Zhang, L.; Wang, B.; Fang, B.; Ma, H.; Yang, Z. et al. (2018): Two-stage short-term wind speed prediction based on LSTM-LSSVM-CFA. *2nd IEEE Conference on Energy Internet and Energy System Integration (EI2)*, pp. 1-6.

Zhang, X.; Sun, L.; Sun, H.; Guo, Q.; Bai, X. (2016): Floating offshore wind turbine reliability analysis based on system grading and dynamic FTA. *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 154, pp. 21-33.

Zhang, Y.; Liu, K.; Qin, L.; An, X. (2016): Deterministic and probabilistic interval prediction for short-term wind power generation based on variational mode decomposition and machine learning methods. *Energy Conversion and Management*, vol. 112, pp. 208-219.

Zhang, Z.; Hong, W. C. (2019): Electric load forecasting by complete ensemble empirical mode decomposition adaptive noise and support vector regression with quantum-based dragonfly algorithm. *Nonlinear Dynamics*, vol. 98, no. 2, pp. 1107-1136.

Zhao, Y.; Ye, L.; Li, Z.; Song, X.; Lang, Y. et al. (2016): A novel bidirectional mechanism based on time series model for wind power forecasting. *Applied Energy*, vol. 177, pp. 793-803.