Australasian Universities Power Engineering Conference, AUPEC 2020, Hobart, TAS, Australia, 29 November – 3 December 2020

Decomposition-Based Short-Term Wind Power Forecasting for Isolated Power Systems

William Aitken
School of Engineering
University of Tasmania
Hobart, Australia
wjaitken@utas.edu.au

Michael Negnevitsky
School of Engineering
University of Tasmania
Hobart, Australia
michael.negnevitsky@utas.edu.au

Evgenii Semshchikov
School of Engineering
University of Tasmania
Hobart, Australia
evgenii.semshchikov@utas.edu.au

***Abstract*—Wind energy penetration has increased significantly and is playing a crucial role in the conversion of power systems to renewable energy. Remote and isolated power systems are increasing wind generation due to high cost of diesel fuel and transportation. To address the concerns of system frequency and scheduling from high penetration of stochastic wind generation, accurate short-term wind power forecasting is required. The research Investigates temporal resolution of wind energy data to improve neural network based forecast models. High resolution wind power data is used to simulate different temporal resolution, for both 10 minute and 1 hour forecast horizons. Three decomposition methods are compared wavelet, empirical mode, and variable mode decomposition. They each decomposed the sampled data into different modes, firstly a long-term component of lower frequencies, then more modes with detailed higher frequency components. To evaluate the temporal resolution and decomposition methods Back propagation neural network (BP), long short-term memory neural network (LSTM) and a convolutional neural network (CNN) are evaluated using wind power data from the King Island power system.**

Keywords—VMD, EMD, WD, LSTM, CNN, BP, temporal resolution, renewable energy.

I. Introduction

Isolated power systems (IPS) provide energy for remote or island communities. Traditional power grids benefit from interconnection and economies of scale that make electricity affordable, isolated communities do not have access to affordable electricity, due to their geographic isolation and size. Conventionally IPS depend on diesel generation, due to low install cost, simplicity and reliability. Though diesel generators have significant operational costs from diesel transportation to remote locations [1] In order to reduce the cost of energy and greenhouse emissions, the use of renewable energy sources (RES) is being integrated. With abundant RE such as wind and solar, the generation is stochastic and intermittent leading to system security and reliability concerns.

One solution is energy storage systems (ESS), which are used for energy shifting, load peak shaving, power quality and spinning reserve. Energy shifting and load peak shaving can have a slower response but requires more storage, suitable ESS are pumped Hydroelectric, hydrogen or compressed air storage. Power quality can be improved by fast responding batteries, super-capacitors or superconducting magnetic energy storage, while spinning reserve can be increased with the installation of flywheels [2]. In Australia IPS have achieved high renewable penetration with the aid of energy storage, King Island was the first to achieve greater than 60% p.a. with multiple ESS's, a lead-acid battery and two flywheels. 60% renewable energy (RE) penetration was also achieved on Flinders island, with reduced battery capacity and four generators of increasing capacity. With low load diesel and demand side management technologies Rottnest island is achieving greater than 50% RE penetration without a battery.

Wind generation is the fastest growing renewable energy source, with the abundance of wind and low operating costs [3]. Wind turbines provide most of the renewable generation in IPS, though the generation is significantly stochastic and intermittent. To effectively manage generation imbalance and control applications of energy storage, accurate short term wind power forecasting is required [4] [5]. Wind power forecasting can be divided into two categories; Uncertain and certain forecasts. Uncertain forecasts are based on probabilistic distributions and interval forecasts, though short term fluctuations make the intervals large for short term forecasting. Certain forecasting can be done by a physical model that uses meteorological data to create generalised mathematical equations, making them complex to implement and site specific. Statistical models can avoid the complexity of grasping the physical mechanism [6], which include support vector machines and artificial neural networks. Due to the characteristics of wind power being complex, Statistical methods are able to learn historic wind power characteristics and then identify them during the forecast process.

There are three artificial neural network models that show promise in the field of wind power forecasting. Back propagation (BP) multi-layer perceptron networks [7] have a long history of use in wind forecasting. The convolutional neural networks (CNN) [8] have been brought from image recognition applications, to provide accurate wind power forecasting. The Long-short term memory (LSTM) is a deep learning network that specialises in time series learning, from speech recognition to wind power forecasting. Additionally in order to improve forecasting accuracy, researchers have decomposed and analysed the wind power signal, mainly by wavelet decomposition (WD), empirical mode decomposition (EMD) and variational mode decomposition (VMD) [9]. A comparison of these techniques will be made in this paper.

II. Decomposition Methods

Decomposition methods convert the wind power series into a set of constitutive series. These series aim to each present a more predictable behaviour than the original signal. They are broken down into approximate and detailed levels of a given signal, where approximation is the general trend of the signal and detail includes the high-frequency components of the signal. The obtained approximations (An) and details (Dn) have better outliers and lower uncertainty, producing signals that are easier to predict for the wind forecasting model.

The decomposition methods are configured to decompose the wind power signal into three decomposition layers/IMF's.

Three layers were shown to have a lower forecast error than five layers for short term wind power forecasting in [6].

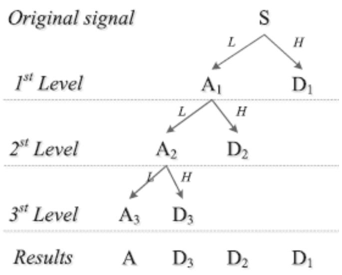

Fig. 1. Decomposition process

A. Wavelet Decomposition

Wavelet decomposition uses the wavelet transform (WT) to decompose a signal. WT's can be divided into two categories: continuous wavelet transform (CWT) and discrete wavelet transform (DWT) [10].

The wavelet transforms scale and translate the mother wavelet to represent the original signal. $W(m,n)$ is the wavelet representation of the signal with respect to the mother wavelet. This requires continuous scaling and translation of the mother signal, creating substantial redundant information. To improve this process the mother wavelet can be scaled and translated using powers of two. This scheme is known as the DWT and is more efficient and just as accurate [11]. Shown in equation 1.

$$W(m,n) = 2^{\frac{m}{2}} \sum_{t=0}^{T-1} f(t)\phi\left(\frac{t-n.2^m}{2^m}\right) \tag{1}$$

The scheme used for wavelet decomposition is a fast DWT algorithm based on filters (deconstruction and reconstruction high and low pass filters) known as the Mallat's algorithm. Multi-resolution via this algorithm is a procedure to obtain "approximations" and "details" from a given signal. Multi-level decomposition can be achieved by successive decomposition of the approximate signals until the desired level is reached. In this paper the 1st Coiflets function is chosen to be the mother wavelet, due to its superior performance over Daubechies functions in the testing conducted.

B. Empirical Mode Decomposition

Empirical mode decomposition (EMD) was proposed by Norden E. Huang when he was also pioneering Intrinsic mode functions (IMF). It produces a unique signal decomposition with good local characteristics in time and frequency domain. EMD can be expressed as equation 2, where the different IMF's $C_i(t)$ when added together with the residual $R_n(t)$equal the original wind power signal $Y(t)$ [12].

$$Y(t) = \sum_{i=1}^{n} C_i(t) + R_n(t) \tag{2}$$

The IMF's are derived through a sifting process; All local maxima are connected by a cubic spline known as the upper envelope, all local minima are also connected by a cubic spline known as the lower envelope. Then the mean envelope $M(t)$ is calculated. The mean envelope is extracted from the original signal by equation 3. Leaving the details $Z(t)$. The process is repeated until $Z(t)$meets the

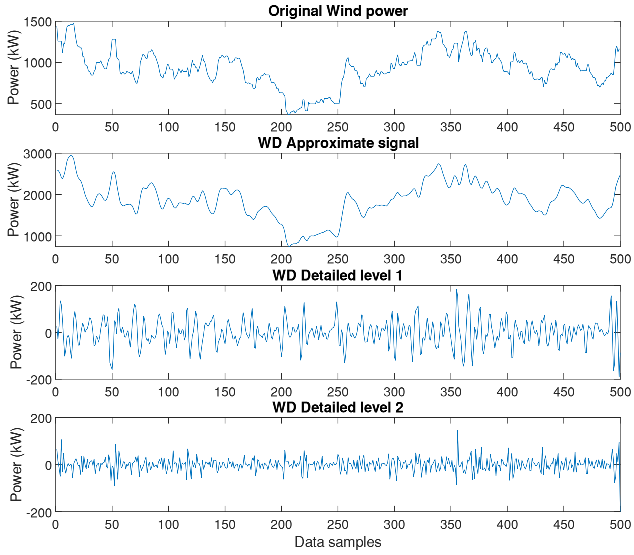

Fig. 2. Wavelet decomposition

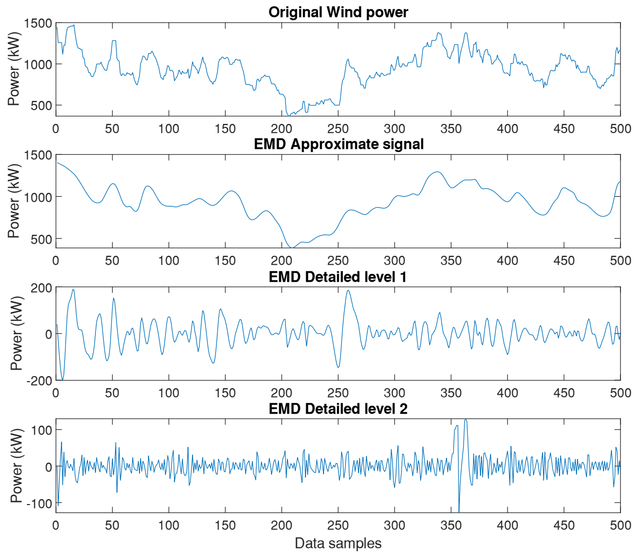

Fig. 3. Empirical mode decomposition

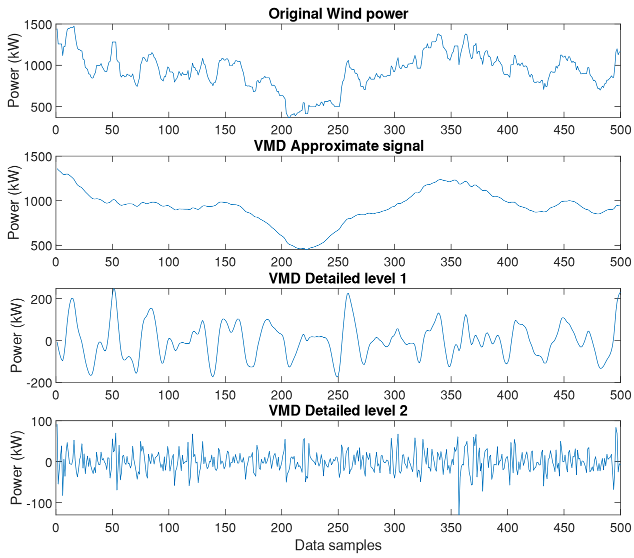

Fig. 4. Variational mode decomposition

properties of a IMF or the maximum decomposition level is reached.

$$Z(t) = Y(t) - M(t) \tag{3}$$

C. *Variational Mode Decomposition*

A new non-stationary and adaptive decomposition method known as variational mode decomposition (VMD) was proposed by the authors of [13]. VMD is able to decompose a real value input signal into a discrete number of sub-signals (modes) that have specific sparsity properties while being able to reproduce the input. The steps of VMD are:

1) *For each mode, the Hilbert transform is used to obtain a unilateral frequency spectrum.*

2) *The mode's frequency spectrum is shifted to "baseband", by mixing with an exponential that is tuned to the respective estimated centre frequency.*

3) *The bandwidth is estimated through the H^1 Gaussian smoothness of the demodulated signal.*

A comparison between the decomposition methods is shown in Fig. 3, Fig. 4 and Fig. 5. The approximate signal in the WD contains more detail, while the EMD and VMD provide a more approximate signal. This same characteristic follows into the first detailed IMF, where the WD is aiming to produce signals similar to wavelets, where as EMD and VMD are trying to isolate certain characteristics.

III. FORECASTING MODELS

Artificial neural networks (ANN) are a popular supervised machine learning approach to wind power forecasting. They are inspired by biological neural networks. Consisting of different layers with neurons that have weighted connections, allowing it to learn non-linear relationships.

A. *Multi-layer Perceptron with Back Propagation*

Multi-layer perceptron (MLP) using Back-propagation (BP) is one of the most commonly used ANN and have been widely studied in the prediction of wind power and wind speed [7]. It is able to realize any non-linear signal with a simple 3-layer structure (input, hidden and output layer).

The BP-MLP model is implemented as a hybrid model by first decomposing the wind power signal into three modes. The three decomposed signals are then used to train three separate BP neural networks. Each network containing 30 neurons in the hidden layer as this was found to perform optimally for the wind power data. After each separate mode is forecast it is recombined to form the forecast signal.

B. *Long Short Term Memory*

The long short-term memory (LSTM) network was chosen as a wind power forecasting model, because of its promising results in recent literature [6]. The LSTM is a recurrent neural network (RNN), it is similar to the MLP neural network, though as it learns information from the last forecast is passed through to the next. Traditional RNN have the problem of vanishing gradients, which hampers the learning of long data sequences. The LSTM was created to counteract this problem, by introducing a forget gate. Allowing the network to be recurrent but also handle long data sequences.

The LSTM was configured with 100 hidden neurons, though it can produce good results with less than 50 neurons in the hidden layer. The 100 neurons produced the best results, with the improvements in higher frequency components of the signal. There is also two ways to implement RNN, either sequence or recurrent. The sequence method has an output every time the network has inputs, whereas the recurrent implementation only outputs at the end of the data sequence. The recurrent implementation was found to be more computationally efficient and also produced better results, as each decomposed signal was able to be input into a single network one after the other, then have a reconstructed output. This removed the reconstruction step as the network was able to learn to reconstruct the original signal during training.

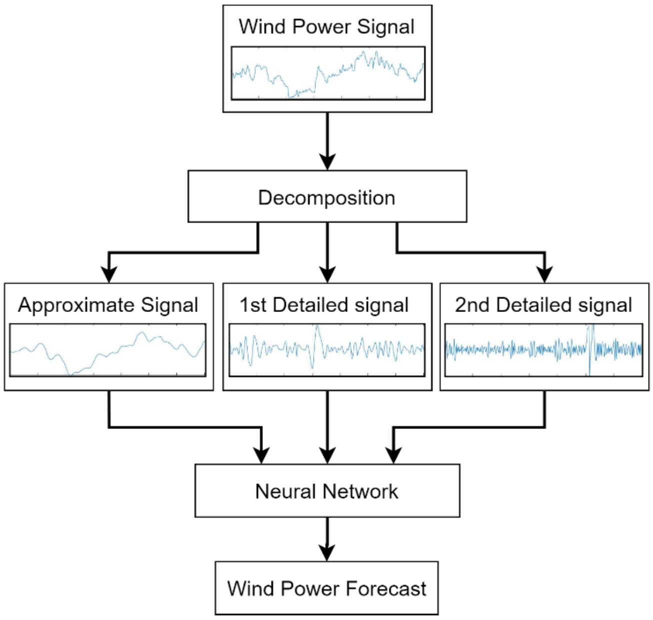

Fig. 5. Flowchart of decomposition and forecasting neural network

C. *Convolutional Neural Network*

The convolutional neural network (CNN) was first proposed by LeChun et al in 1998 and has been widely applied in classification of images. The CNN model has also been shown to achieve good forecast performance of wind speed. The structure of a CNN is essentially a MLP network, where the input is a 2D image that is filtered reducing over-fitting problems, allowing parallelism and fast training. A typical CNN consists of an input layer, convolution layer, pooling layer, fully-connected layer and regression layer.

The inputs to the model were arranged in a 2D matrix with the three decomposition signals along the columns and the inputs down the rows. This created a single NN that could reconstruct the original signal similar to the LSTM model. In the convolution layer 100 filters was used.

IV. FORECAST HORIZONS

The majority of the literature focuses on wind forecasting for large grid connected wind generation. In deregulated energy markets generation units are allocated in one day-ahead schedules for operational security, and incremental generation changes are made hourly [14]. For wind generation to commit to these requirements they require accurate one hour and 24 hour forecasts. The models proposed for this will either use multiple models to forecast one hour and 24 hours [15] or use multi-step forecasting for every hour over the 24 hours [16].

Isolated power systems are not operated as a market, but as a single hybrid power system. The high penetration of renewable energy and small size of the system creates grid stability problems. Requiring shorter forecast horizons for load and generation, which is primarily wind power. The hierarchy control system of an isolated power system has four stages: Inertial, Primary, secondary and tertiary frequency response. The inertial response is almost instantaneous and the primary response regulates voltage and frequency within seconds of it changing. The offset left by primary control can be removed by adjusting dispatchable assets with secondary control, which operates from 30s to 15min. Finally tertiary control maintains the optimal operating conditions, by re-scheduling generation using economic or priority dispatch. Scheduling for duration's of 5min to 1hr [2] [17].

Forecasting horizons that will be considered in this paper are:

- *10-minute forecast:* Secondary control: Reserve and frequency control
- *1-hour forecast:* Tertiary control: Unit Commitment and economic dispatch

V. CASE STUDIES

A. Error Indicators

The error indicators currently used to evaluate wind power forecasting are mean absolute percentage error (MAPE), root mean squared error (RMSE) and normalised RMSE (NRMSE).

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\frac{P_t^w - P_t^f}{P^w} \times 100 \tag{4}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(P_t^w - P_t^f\right)^2} \tag{5}$$

Where P_t^w and P_t^f are measured and forecast wind power at sample t, respectively. N is the number of data samples tested. P_w is the mean of the sample data and P_c is the capacity of the wind farm.

Both MAPE and NRMSE are relative error calculations and can be used to compare accuracy of various forecasting methods. Though NRMSE is not a reliable error calculation when used to compare forecast error of different wind farms. The power output can vary between zero and the wind farm capacity, where some wind farms may average 50% of capacity and others 70%. Therefore, NRMSE is not used to evaluate model performance.

The standard error calculation of RMSE, cannot be used to evaluate model performance on different wind farms. Though is an important measure for power system operation, as it indicates the expected power difference between the forecast and the actual wind power.

B. Data Source

To train and evaluate the wind power forecasting models proposed in this paper, wind power data was gathered from Hydro Tasmania's isolated power system on King island (KI). Operating five turbines the installed wind capacity is 2.45MW, capable of providing up to 100% of the islands energy demands with enabling technology. The temporal resolution of the data is 3 seconds, providing considerable detail that allows for different temporal resolutions to be evaluated. The training samples used are from 1st January 2019 to 31st August 2019. The data from 1st September to 31st December is used to test the forecast models.

C. Analysis of Temporal Resolution

The wind power data from KI represents the instantaneous power output of the wind farm every three seconds. To simulated different temporal resolutions and reduce external errors, the same wind data is used but sampled at different rates. The sampled data is then decomposed by one of the three methods and each mode is normalised. The eight months of training data will have a different data length depending on the temporal resolution. To include the range of unique power characteristics over that time, fifty thousand samples are taken from the decomposed data, that are spread evenly across it. The thousands of samples are then used to train the neural network forecasting models. The testing of the models is done in the same way, to test across the whole four months but also test 25 000 data points.

The two cases that were consider for analysing temporal resolution are the forecast horizons ten minutes and one hour. The forecast errors of the ten minute and one hour forecast are shown in Table I and Table II respectively.

Table I:RMSE (kW) of 10-minute forecast, comparison of models with different data resolutions

Forecast Models	**Temporal Resolution (min)**				
	10	***5***	***2.5***	***1***	***0.5***
LSTM-WD	26.28	**20.22**	71.02	190.30	212.37
LSTM-VMD	149.10	86.43	131.41	193.12	209.55
LSTM-EMD	136.81	148.93	171.03	199.77	211.22
LSTM	183.70	181.80	182.25	180.31	181.17
BP-WD	**136.04**	158.96	181.10	195.74	208.08
BP-VMD	138.63	184.99	243.81	383.88	273.49
BP-EMD	144.51	150.98	166.34	200.91	215.15
BP	220.47	212.18	188.45	290.08	295.52
CNN-WD	345.11	210.79	297.76	241.76	224.10
CNN-VMD	291.17	**180.90**	454.97	200.11	356.25
CNN-EMD	223.00	221.20	384.95	256.75	290.52
CNN	283.45	423.14	279.14	435.32	1470.79

The temporal resolution for a ten minute forecast is compared in Table I , where the lowest error for each NN is highlighted in bold. It shows that when forecasting ten minutes a five minute data resolution has the lowest RMSE for the LSTM and CNN, while a ten minute data resolution has the lowest error for the BP model. For the LSTM and the BP the WD consistently had the lowest error for all temporal resolutions, whereas the CNN error isn't related to the decomposition method. Another interesting thing to note is that the LSTM model with no decomposition had the same error for all data resolution, which was not the case for the other two networks.

Forecasting 1 hour is shown in Table II, where the results are similar. The lowest errors were found for the BP and CNN with a resolution of 30 minutes, the LSTM produced the lowest RMSE with a data resolution of 60 minutes. The WD consistently produced the lowest forecast error for each model.

Table II: RMSE (kW) of 1-hour forecast, comparison of models with different data resolutions

Forecast Models	Temporal Resolution (min)				
	60	*30*	*20*	*10*	*5*
LSTM-WD	**112.97**	165.04	166.14	221.33	276.57
LSTM-VMD	308.90	506.88	291.58	244.69	278.66
LSTM-EMD	242.39	260.71	255.96	262.42	288.59
LSTM	288.35	265.47	267.97	261.53	265.84
BP-WD	205.92	**175.91**	190.50	256.67	264.77
BP-VMD	265.61	604.57	231.08	265.95	273.75
BP-EMD	234.73	198.05	202.33	219.09	264.25
BP	318.98	298.39	299.98	295.59	312.13
CNN-WD	212.63	**157.38**	231.51	301.86	295.00
CNN-VMD	294.35	380.87	254.81	268.32	287.00
CNN-EMD	293.55	243.73	242.34	282.36	328.51
CNN	422.30	340.94	354.86	311.96	295.59

It is shown that for both forecast horizons the LSTM-WD forecasting model consistently produced the most accurate wind power forecast. The other finding is that the optimal temporal resolution is relative to the forecasting horizon, it is either the same or half of the horizon. The detailed parts of the wind data oscillate rapidly and by sampling the data twice for every forecast horizon it creates a signal that has a period equal to the forecast horizon.

D. *Analysis of Input Length*

Table III: RMSE (kW) of 10-minute forecast, comparison of models with different input length

Forecast Models	Input (hrs)				
	0.5	*1*	*2*	*4*	*6*
LSTM-WD	**17.46**	20.30	25.24	35.39	39.75
LSTM-VMD	87.82	86.86	86.02	89.46	98.33
LSTM-EMD	149.66	149.59	153.12	156.40	161.90
LSTM	211.23	210.96	211.23	211.61	214.65
BP-WD	155.25	155.86	156.93	160.94	165.62
BP-VMD	181.11	191.25	202.55	204.68	222.22
BP-EMD	**138.77**	139.55	149.04	167.88	175.22
BP	208.56	209.51	228.48	228.49	262.88
CNN-WD	**145.10**	346.93	249.72	239.14	353.49
CNN-VMD	216.94	224.53	244.41	299.56	253.19
CNN-EMD	193.59	347.25	394.58	283.31	288.93
CNN	284.84	275.98	385.07	243.63	477.43

The optimal temporal resolution was found to be half the forecast horizon. Using the optimal resolution, the models are compared with different input lengths. For both forecast horizons the input length is represented in hours, although temporal resolution of 5 minutes will have 12 data points for every hour and 30 minute data resolution will have 2. It is represented in hours to show how much of the long term trend is required for different forecasting horizons. Table III shows the 10 minute forecast and Table IV shows the results of 1 hour forecast.

The results show that a 10 minute forecast requires only 30 minutes of inputs, equating to six data points. Longer inputs are shown to increase the error of the forecast. This shows that the short horizon requires only the short term details of the signal and the long term trend is less important causing the model to lose focus on the short term detail. For the 1 hour forecast the results show that the lowest error is found with an input length of six hours, equating to 12 data points. This shows that longer horizons require more long-term information of the signal, although when the input is too long it can cause the model to miss detail.

Table IV: RMSE (kW) of 1-hour forecast, comparison of models with different input length

Forecast Models	Input (hrs)				
	2	*6*	*12*	*24*	*48*
LSTM-WD	154.40	**66.52**	76.80	107.08	171.09
LSTM-VMD	443.50	418.01	404.76	393.43	362.02
LSTM-EMD	259.26	251.19	253.98	273.31	296.00
LSTM	320.33	318.56	322.41	325.52	347.94
BP-WD	162.66	153.91	**146.93**	156.95	194.44
BP-VMD	1014.7	861.58	733.34	782.42	762.89
BP-EMD	191.04	205.00	209.31	226.53	254.70
BP	295.69	301.62	318.50	325.95	358.16
CNN-WD	215.21	**198.93**	212.75	280.07	335.14
CNN-VMD	467.15	395.42	874.73	669.90	663.66
CNN-EMD	348.53	293.75	357.31	334.19	546.48
CNN	388.71	398.15	385.98	410.60	395.13

E. *Evaluation of models*

To evaluate the models the best configuration of each neural network model is selected for each forecast horizon. They are compared against the most widely used short-term method of forecasting, the Persistence method. A short four hours portion of the forecast is shown in Fig. 6 and 24 hours in Fig. 7 to compare the forecast of the 10-minute and 1-hour horizon, respectively. Details of the forecast error are shown in Table V and VI.

Table V: Forecast error and calculation time of 10-minute forecast models

Forecast Model	RMSE (kW)	MAPE (%)	Time (s)
LSTM-WD	17.46	0.99	5.02
BP-WD	109.95	7.99	15.87
CNN-WD	145.11	11.40	4.13
Persistence	235.30	14.62	0.01

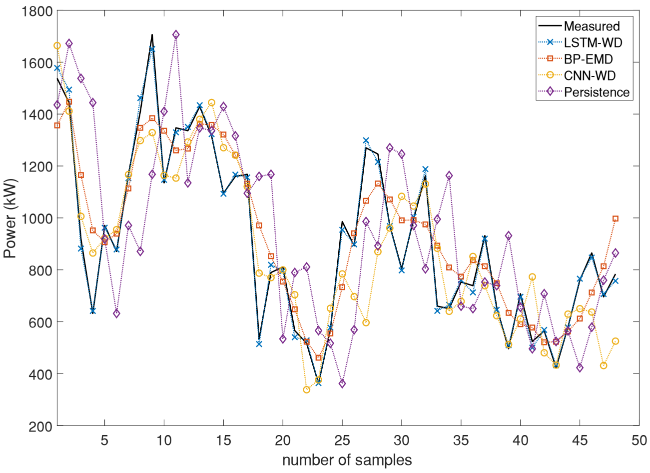

Fig. 6. 10-minute forecast model comparison over four hours

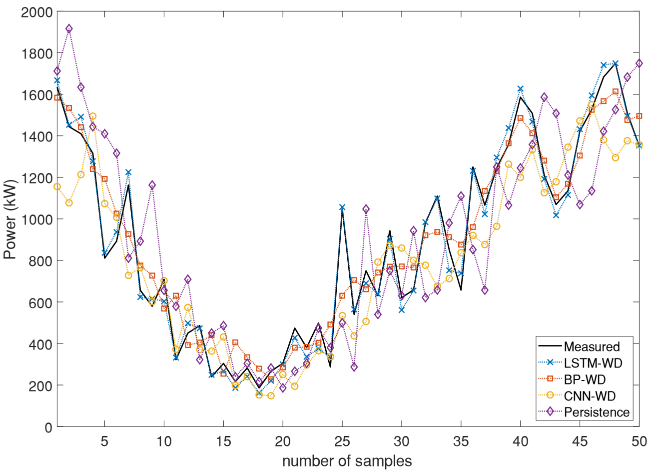

Fig. 7. 1-hour forecast model comparison over 24 hours

Table VI: Forecast error and calculation time of 1-hour forecast models

Forecast Model	**RMSE (kW)**	**MAPE (%)**	**Time (s)**
LSTM-WD	66.52	4.41	4.08
BP-WD	146.95	11.17	3.61
CNN-EMD	293.73	23.88	4.20
Persistence	312.40	24.08	0.01

VI. Conclusions

This paper compared the wind power forecasting of hybrid neural network models using signal decomposition techniques. Three-layer decomposition represented a long-term component, the fluctuation and random characteristics of wind. The WD with the 1st Coiflet as the mother wavelet, produced a signal that was forecast more accurately by the three forecast neural network. The LSTM-WD produced considerably less forecast error than the BP and the CNN, over both 10 minute and 1 hour forecast horizons. Analysis on the temporal resolution of forecasting data found that, data resolutions of half the forecast horizon were optimal. Providing the amount of detail to forecast accurately, while also reducing the excess random fluctuations in the training data. The length of time-steps required to forecast 10 minute is only 30 minutes of data points, and as the forecast horizon increases the model requires longer input with less resolution.

References

[1] D. Nikolic, M. Negnevitsky, and M. de Groot, “Effect of the diesel engine delay on stability of isolated power systems with high levels of renewable energy penetration,” in 2015 International Symposium on Smart Electric Distribution Systems and Technologies (EDST), pp. 70–73, 2015.

[2] E. Semshchikov, M. Negnevitsky, J. M. Hamilton, and X. Wang, “Costefficient strategy for high renewable energy penetration in isolated power systems,” IEEE Transactions on Power Systems, pp. 1–1, 2020.

[3] Z. Shi, H. Liang, and V. Dinavahi, “Direct interval forecast of uncertain wind power based on recurrent neural networks,” IEEE Transactions on Sustainable Energy, vol. 9, pp. 1177–1187, July 2018.

[4] M. Cui, D. Ke, Y. Sun, D. Gan, J. Zhang, and B. Hodge, “Wind power ramp event forecasting using a stochastic scenario generation method,” IEEE Transactions on Sustainable Energy, vol. 6, pp. 422–433, April 2015.

[5] C. W. Potter and M. Negnevitsky, “Very short-term wind forecasting for tasmanian power generation,” IEEE Transactions on Power Systems, vol. 21, no. 2, pp. 965–972, 2006.

[6] L. Han, R. Zhang, X. Wang, A. Bao, and H. Jing, “Multi-step wind power forecast based on vmd-lstm,” IET Renewable Power Generation, vol. 13, no. 10, pp. 1690–1700, 2019.

[7] Gang Zhang, Lei Zhang, and Tuo Xie, “Prediction of short-term wind power in wind power plant based on bp-ann,” in 2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), pp. 75–79, 2016.

[8] M. Solas, N. Cepeda, and J. L. Viegas, “Convolutional neural network for short-term wind power forecasting,” in 2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), pp. 1–5, 2019.

[9] Z. Qian, Y. Pei, H. Zareipour, and N. Chen, “A review and discussion of decomposition-based hybrid models for wind energy forecasting applications,” Applied Energy, vol. 235, pp.939–953, 02 2019.

[10] C. Lei and L. Ran, “Short-term wind speed forecasting model for wind farm based on wavelet decomposition,” in 2008 Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies, pp. 2525–2529, 2008.

[11] H. zhi Wang, G. qiang Li, G. bin Wang, J. chun Peng, H. Jiang, and Y. tao Liu, “Deep learning based ensemble approach for probabilistic wind power forecasting,” Applied Energy, vol. 188, pp. 56 – 70, 2017.

[12] E. Dokur, M. Kurban, and S. Ceyhan, “Hybrid model for short term wind speed forecasting using empirical mode decomposition and artificial neural network,” in 2015 9th International Conference on Electrical and Electronics Engineering (ELECO), pp. 420–423, 2015.

[13] K. Dragomiretskiy and D. Zosso, “Variational mode decomposition,” IEEE Transactions on Signal Processing, vol. 62, no. 3, pp. 531–544, 2014.

[14] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal, “A review of wind power and wind speed forecasting methods with different time horizons,” in North American Power Symposium 2010, pp. 1–8, 2010.

[15] O. M´egel, J. L. Mathieu, and G. Andersson, “Scheduling distributed energy storage units to provide multiple services under forecast error,” International Journal of Electrical Power & Energy Systems, vol. 72, pp. 48 – 57, 2015. The Special Issue for 18th Power Systems Computation Conference.

[16] V. Kolev and S. Sulakov, “Forecasting the hourly power output of wind farms for day-ahead and intraday markets,” in 2018 10th Electrical Engineering Faculty Conference (BulEF), pp. 1 -4, 2018.

[17] R. Doherty and M. O’Malley, “A new approach to quantify reserve demand in systems with significant installed wind capacity,” IEEE Transactions on Power Systems, vol. 20, no. 2, pp. 587–595,