



Article

A Combined Fuzzy GMDH Neural Network and Grey Wolf Optimization Application for Wind Turbine Power Production Forecasting Considering SCADA Data

Azim Heydari ^{1,*}, Meysam Majidi Nezhad ¹, Mehdi Neshat ², Davide Astiaso Garcia ³, Farshid Keynia ⁴, Livio De Santoli ¹ and Lina Bertling Tjernberg ⁵

- ¹ Department of Astronautical, Electrical and Energy Engineering (DIAEE), Sapienza University, 00184 Rome, Italy; meysam.majidinezhad@uniroma1.it (M.M.N.); livio.desantoli@uniroma1.it (L.D.S.)
- ² Optimization and Logistics Group, School of Computer Science, University of Adelaide, Adelaide 5005, Australia; neshat.mehdi@gmail.com
- ³ Department of Planning, Design, and Technology of Architecture, Sapienza University of Rome, 00197 Rome, Italy; davide.astiasogarcia@uniroma1.it
- ⁴ Department of Energy Management and Optimization, Institute of Science and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman 7631133131, Iran; fkeynia@gmail.com
- ⁵ School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology Stockholm, 10044 Stockholm, Sweden; linab@kth.se
- * Correspondence: azim.heydari@uniroma1.it



Citation: Heydari, A.; Majidi Nezhad, M.; Neshat, M.; Garcia, D.A.; Keynia, F.; De Santoli, L.; Bertling Tjernberg, L. A Combined Fuzzy GMDH Neural Network and Grey Wolf Optimization Application for Wind Turbine Power Production Forecasting Considering SCADA Data. *Energies* **2021**, *14*, 3459. <https://doi.org/10.3390/en14123459>

Academic Editor:
Francesco Castellani

Received: 7 May 2021
Accepted: 9 June 2021
Published: 11 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: A cost-effective and efficient wind energy production trend leads to larger wind turbine generators and drive for more advanced forecast models to increase their accuracy. This paper proposes a combined forecasting model that consists of empirical mode decomposition, fuzzy group method of data handling neural network, and grey wolf optimization algorithm. A combined K-means and identifying density-based local outliers is applied to detect and clean the outliers of the raw supervisory control and data acquisition data in the proposed forecasting model. Moreover, the empirical mode decomposition is employed to decompose signals and pre-processing data. The fuzzy GMDH neural network is a forecaster engine to estimate the future amount of wind turbines energy production, where the grey wolf optimization is used to optimize the fuzzy GMDH neural network parameters in order to achieve a lower forecasting error. Moreover, the model has been applied using actual data from a pilot onshore wind farm in Sweden. The obtained results indicate that the proposed model has a higher accuracy than others in the literature and provides single and combined forecasting models in different time-steps ahead and seasons.

Keywords: power system; wind power production; SCADA data; fuzzy GMDH neural network; grey wolf optimization

1. Introduction

Wind power industries have been tremendously expanded and are expected to progress at a compound annual growth rate (CAGR) of 5.2% between 2020 and 2027. This extension resulted in the produced power cost of wind energy as one of the most significant renewable and low-carbon energy resources. Wind power generation is currently one of the principal renewable energy power generations [1–4]. Wind energy is stochastic, uncertain, and discontinuous, antagonistically influencing the power grid's protected and stable activity and the nature of the power supply [5]. The stochasticity and discontinuity of wind power could diminish the reliability prediction system and wind power quality [6].

A potential answer to these issues is to improve the forecast accuracy of wind generation. Several studies [7–12] are proposed to portray the distribution of the wind power prediction, and diverse scientific methodologies are connected to improve its accuracy.

Other studies proposed complex models such as the Laplace distribution [9], the Beta distribution [10], the hyperbolic distribution [11], the Levy α -stable distribution [12], and the flexible likelihood distribution [13], which have been proposed to improve the fitting precision of the wind power prediction. In the previous decade, many studies to assess and predict the various aspects of energy management and power systems have been presented. For example, wind and solar power generation forecasting [14–16], condition monitoring of wind turbines [17], electricity market [18], and load forecasting [19,20] are proposed. Amjady et al. (2011) provided the short-term wind power prediction dependent on the ridgelet neural network (RNN) with a high capacity estimate ability. They suggested a differential evolution algorithm with a new selection mechanism and crossover to train the network [21]. Han et al. (2017) proposed combined models based on autoregressive integrated moving average (ARMA) and non-parametric model for wind speed forecasting [22].

The results demonstrated that non-parametric based combined models usually have a better performance than other models. Jonas C. Pelajo et al. (2019) developed a model to predict wind speed and energy price to determine the optimal maintenance planning of a real wind farm in the Brazilian Northeast [23]. Osório et al. (2014) proposed a combined forecasting model based on mutual information, wavelet transform, particle swarm optimization, and adaptive neuro-fuzzy inference system framework to predict the short-term wind power and electricity market prices [24]. Gallego-Castillo et al. (2016) provided a quantile relapse model dependent on the recreating piece Hilbert space (RKHS) system to predict the wind power probabilistic. Furthermore, they implemented two types of models (online and offline) for a real wind farm [25]. Xiao et al. (2017) employed an electrical power system prediction model using a wavelet neural network (WNN) model and an improved cuckoo search algorithm. The results showed that the proposed model essentially diminished the expectation error with respect to other relative models [26].

Kunpeng Shi et al. (2018) provided a combined model based on two-stage feature selection and improved random forest models to short term wind power forecasting [27]. Van Quang Doan et al. (2019) have presented a mesoscale ensemble model to predict wind speed ramps. The proposed model applied at real wind farms in Japan [28]. Duan et al. (2021) developed a combined intelligent model based on the improved variational mode decomposition and Correntropy long short-term memory neural network to predict wind power. The model was evaluated using two wind farms in China at different sampling intervals [29]. Yildiz et al. (2021) presented a two-step new deep learning approach based on the variational mode decomposition (VMD) method and modified the residual-based deep convolutional neural network for wind power forecasting [30]. Jafarzadeh et al. (2021) provided a modified fuzzy wavelet neural network for short-term wind power forecasting considering weather and power plant parameters. In order to evaluate the model, the Mnjil wind power plant in Iran has been used [31].

In addition, GIS-based models play an important role in renewable energy potential assessment and prediction [32,33]. Furthermore, the behavior and performance of renewable energy systems can be estimated using GIS models [34,35].

Generally, in order to model the wind turbine power production analysis, a combined intelligent solution is required. It means that the data should first be modelled by a combined data pre-processing model then a combined intelligent strategy should analyze the processed data. This type of strategy plays an essential role in managing the energy production of wind farms.

In this research, we propose an integrated strategy that couples an empirical mode decomposition, fuzzy GMDH (group method of data handling) neural network, and grey wolf optimization algorithm (GWO) to forecast the produced power of wind turbines. Furthermore, to detect and clean outliers, a combined K-means and density-based local outliers (LOF) are applied.

The main contributions and novelty of this paper are illustrated as follows:

- (a) K-means is one of the most well-known clustering methods, a fast and efficient technique in unsupervised learning methods. However, it suffers from some deficiencies such as (i) predefining the number of clusters and centers in advance, (ii) not being able to handle noisy data and outliers properly, and finally, k-means is not proper to classify clusters with non-convex shapes. In order to deal with these listed issues, we proposed a combination of identifying density-based local outliers (LOF) and k-means for cleaning the raw supervisory control and data acquisition (SCADA) data as the initial section of the pre-processing.
- (b) As wind speed and power forecasting involve the non-linear power curve, stochastic and noisy behavior of the recorded wind data, the empirical mode decomposition (EMD) method is proposed to deal with these uncertainties and increase the modelling accuracy.
- (c) Fuzzy-GMDH neural networks are considered as one of the most effective methods to model the time-series data with high-level noise and short input sampling. However, initializing the hyper-parameters of fuzzy-GMDH is challenging and time-consuming. With regards to adjusting the hyper-parameters, we apply a robust and fast search method called the grey wolf optimization (GWO) algorithm. Applying the GWO as a hyper-parameter tuner improved the proposed model's accuracy and reliability to forecast wind power.
- (d) The proposed combined forecasting model has successfully verified on two actual wind turbines SCADA datasets. In addition, the proposed forecasting model is compared with the other valid combined forecasting models.

2. Materials and Methods

After a brief description of the SCADA system and data gathering, this section illustrates the artificial intelligence methods proposed in this paper.

2.1. SCADA System

The SCADA system, known as remote supervision and control of wind turbines in wind farms, plays a significant role in the wind power forecasting models. This paper's collected and applied SCADA data is related to a large wind farm (located in Sweden). The input data includes the power output of wind turbines and wind speed (short-term with the interval of 10 min) for a year from Jan to Dec 2015. Furthermore, in order to evaluate and compare the performance of the proposed hybrid model, we applied the SCADA data for two wind turbines (wind turbine 1 (WT1) and wind turbine 2 (WT2)).

2.2. Proposed Wind Power Forecasting Strategy

In this study, a multi-step hybrid intelligent model has been proposed as a means to predict wind power production (see Figure 1).

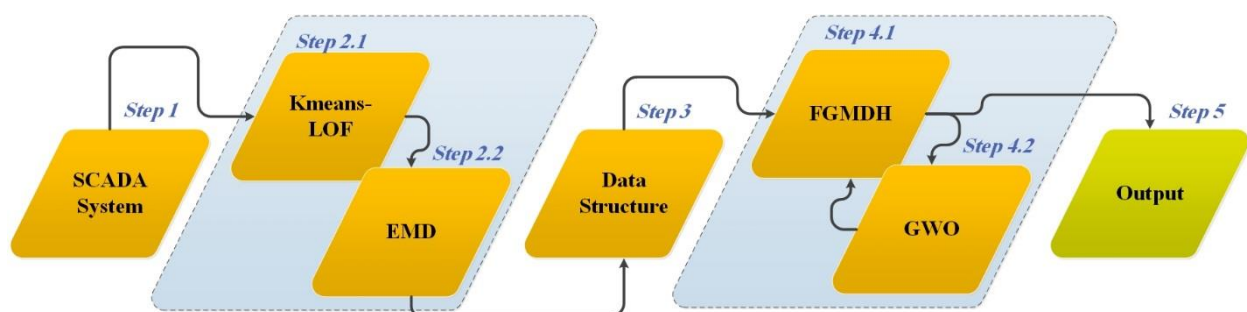


Figure 1. General schematic of the proposed forecasting model.

Due to the wide range of intelligent methods such as neural networks and metaheuristic optimization algorithms, in this paper, we presented a hybrid forecasting model based

on FGMDH and GWO for wind power production forecasting. The GWO optimization algorithm can perform the neural network (FGMDH) training step well and optimize the value of network parameters. Therefore, this algorithm (GWO) is constructive to the performance of the proposed model for predicting wind power production.

Since the structure of the input matrix plays a significant role in determining the output and accuracy of the model, in the first place, the various input signals (wind power and wind speed) are decomposed through the EMD method to different high and low frequencies (see Figure 2).

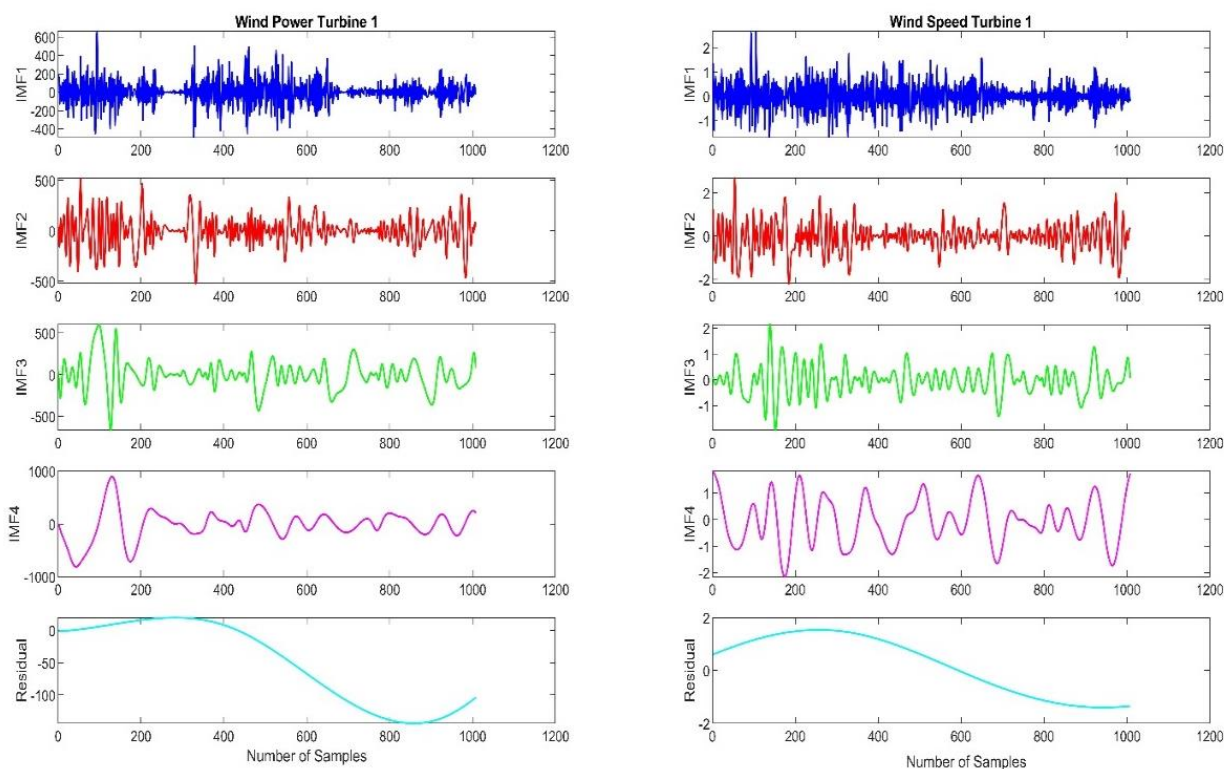


Figure 2. The EMD output, i.e., decomposition signals of wind power and wind speed of turbine 1 (WT1).

Five types of decomposed frequencies (IMF1, IMF2, IMF3, IMF4, residual) are selected and applied by delaying a unit of time ($t-1$) as inputs of the model subsequently (see Figure 3—inputs and output data structure). In addition, the lagged values (1 to 5) for the original wind power signal and actual wind speed signal are considered as input parameters (Figure 3—inputs and output data structure). In the next step, the FGMDH method has been employed to predict the wind turbine power.

The FGMDH model structure includes different neurons. The parameters grouped in the form of Gaussian variables and the weight of the fuzzy rule in each neuron are unknown. In this paper, the GWO algorithm is applied with the purpose of optimizing the FGMDH model variables (the group-unknown variables in neurons).

In this study, in order to evaluate the performance and reliability of forecasting models, the wind turbine power production is predicted for different seasons at two times (10-min and 1-h). The framework of the proposed model is represented in Figure 3.

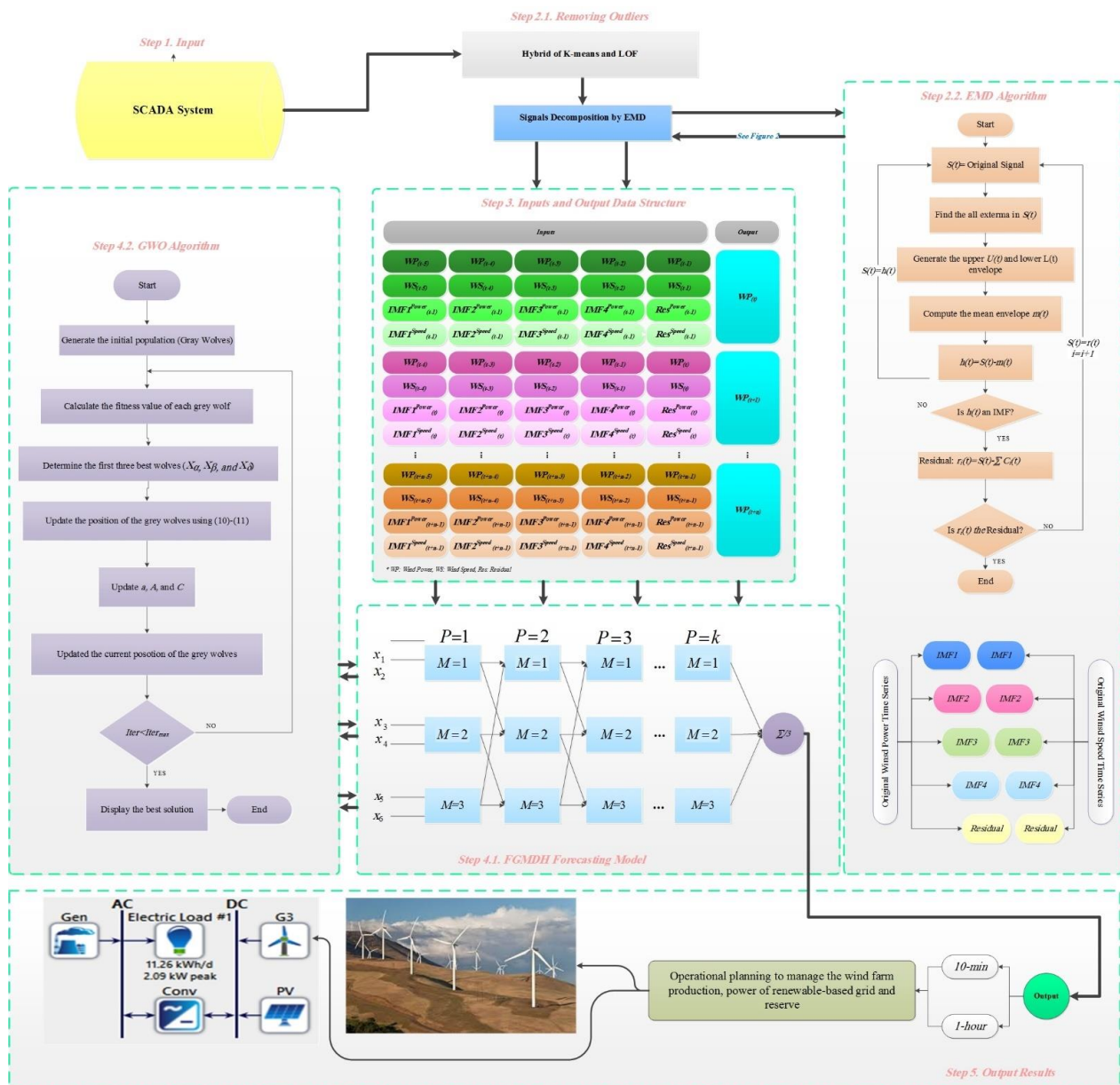


Figure 3. The framework of the proposed model.

2.3. Data Cleaning

The raw SCADA datasets usually include different forms of noise that directly negatively affect the accuracy of the forecasting process. One of the most notable outliers can be the negative wind turbine power outputs observed when the wind speed is shallow or during a failure situation. For evaluating the distribution of the raw SCADA dataset, Figure 1 is plotted, and also an abnormal distribution of wind power can be seen in Figure 4.

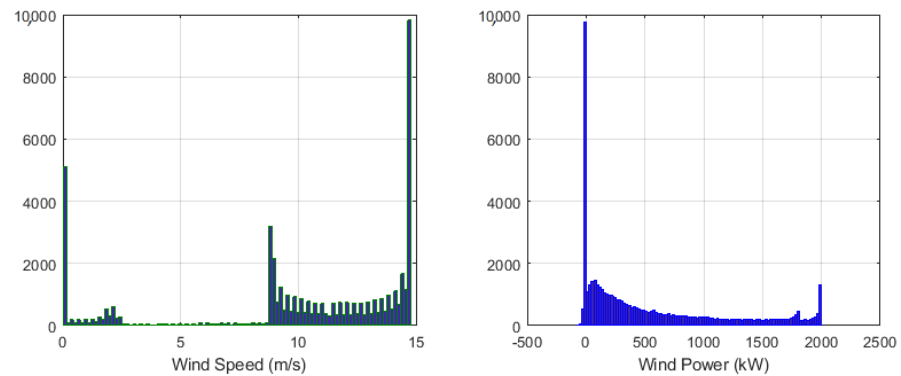


Figure 4. The distribution of the raw SCADA data WT1 (wind speed and power).

In the pre-processing section, it is recommended that [36] these negative powers should be set as zero. In addition, to remove the impact of the data scale, a Min-Max normalization is implemented for the feature scaling. Meanwhile, as each wind turbine has a unique power curve that presents the average efficiency of the applied wind turbine, without declaring the particular mechanical components, Figure 5 is plotted for showing this characteristic of the first wind turbine in this research. The scatter data point indicates the outliers.

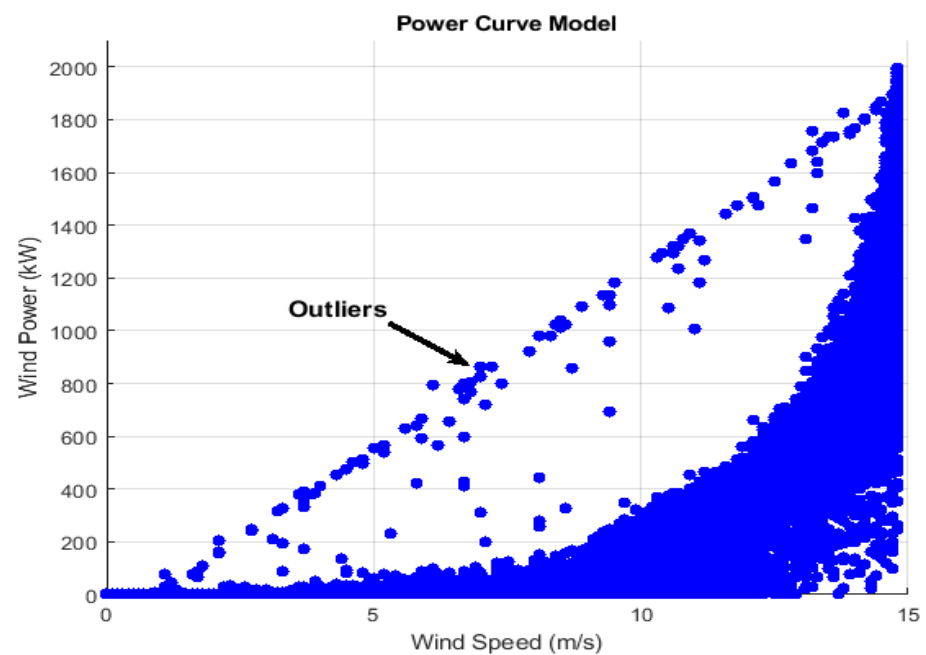


Figure 5. The power curve model of the WT1.

The proposed cleaning data method is a combined K-means clustering and the identifying density-based local outliers (LOF) method [37]. In the first step, a k-means clustering method is employed to classify the SCADA data into various clusters. Then, in each cluster, the local density-based method is adopted to eliminate the potential noises. The clean data after using K-means clustering and the LOF method can be illustrated in Figure 6.

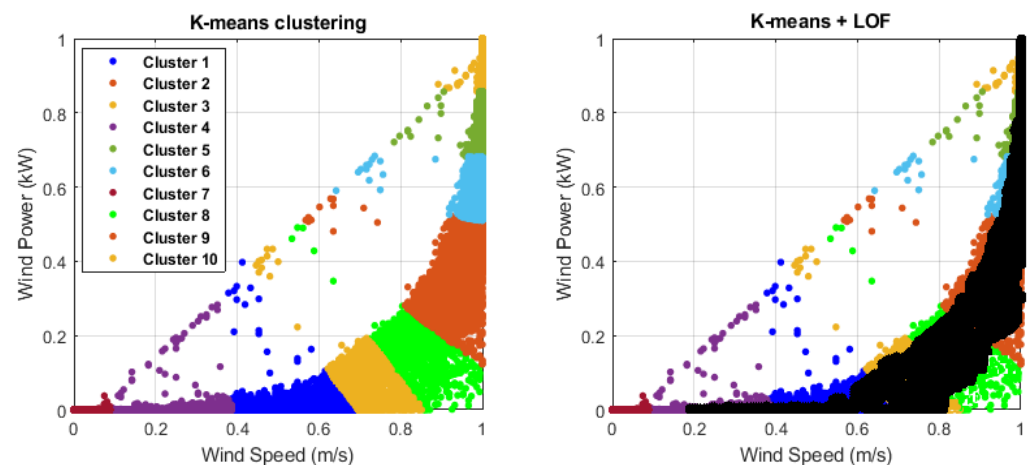


Figure 6. The K-means clustering method performance and the applied data (WT1) divided into 10 clusters (left). The dark points represent the clean SCADA data after applying the LOF (right).

2.4. Empirical Mode Decomposition

The EMD is a method of signal decomposition that can analyse the non-linear and non-stationary time series. Moreover, using this method is more accessible and more understandable compared to wavelet decomposition [38,39]. In addition, EMD does not stand in the need of deciding a mother function in advance (beforehand of time) by no means such as wavelet decomposition. The most important characteristic of the EMD is a fully data-driven decomposing means by which signals break down into various independent components within the interval of local specifications of a signal.

Decomposing initial signals as intrinsic mode functions (IMFs) and residual into a finite amount of oscillatory functions is the concept of EMD. These IMFs must be met by the following conditions: (1) The number of extreme must be equal to the number of zero crossings or their maximum difference is equal to one; (2) the mean value of the envelopes characterized by local maxima and local minima must be zero at all components.

The EMD is a sifting method using a real signal to extract the IMFs and residual. The calculation of the EMD can be given as the following steps [38,39]:

Stage 1: Recognize all local maxima and local minima in time series $S(t)$.

Stage 2: Connect all local maxima and minima to produce the upper $U(t)$ and lower $L(t)$ envelopes using a cubic spline line.

Stage 3: Calculate the point-by-point mean envelope from the upper and lower envelopes and create the mean envelopes $m(t)$ later as:

$$m(t) = \frac{[U(t) + L(t)]}{2} \quad (1)$$

Stage 4: Compute the distinction between the mean envelopes and the actual signal:

$$h(t) = S(t) - m(t) \quad (2)$$

Stage 5: Check whether $h(t)$ is an intrinsic mode function (IMF). Provided that this is true, it is treated as the i th IMF and afterwards the actual time series is supplanted by the residuals $h(t) = S(t) - m(t)$. If not, is supplanted by $h(t)$.

Stage 6: Repeat Steps 1–5 until the standard deviation magnitude of the two consecutive sifting results (IMFs and Residual) is lower than the predefined stopping criterion.

Using the above-mentioned sifting process, many IMFs can be obtained from high frequency to low frequency, thereby disintegrating into several IMFs and a residual as:

$$r_n(t) = S(t) - \sum_{i=1}^n C_i(t) \quad (3)$$

where $r_n(t)$ and n are the last residuals and the number of IMFs, respectively. $C_i(t)$ ($i = 1, 2, \dots, n$) indicates different IMFs.

2.5. Fuzzy-GMDH Model

The FGMDH is a machine learning strategy in the hierarchical structure [40]. In this model, every neuron has two inputs and an output. The general structure of the FGMDH system was shown in Figure 3 (the FGMDH forecasting model). In this figure, the output of each neuron in each layer is considered as the input in the following layer. The last output is determined to utilize the mean of the last layer output.

The FGMDH structure part in Figure 6 demonstrates that the inputs from the m th model and p th layer are the outputs of the $(m - 1)$ th and m th model in the $(p - 1)$ th layer. The numerical function for computing the y^{pm} (the output variable of the m th model in the p th layer) is as follows:

$$y^{pm} = f(y^{p-1,m-1}, y^{p-1,m}) = \sum_{k=1}^K \mu_k^{pm} \cdot w_k^{pm} \quad (4)$$

$$\mu_k^{pm} = \exp \left\{ - \frac{(y^{p-1,m-1} - a_{k,1}^{pm})^2}{b_{k,1}^{pm}} - \frac{(y^{p-1,m} - a_{k,2}^{pm})^2}{b_{k,2}^{pm}} \right\} \quad (5)$$

where w_k^{pm} and μ_k^{pm} are its corresponding weight parameter and the k th Gaussian function, respectively. Moreover, a_k^{pm} and b_k^{pm} are the Gaussian parameters [41]. Furthermore, the last output is calculated by the following equation:

$$y = \frac{1}{M} \sum_{m=1}^M y^{pm} \quad (6)$$

The learning procedure of feed forward FGMDH is known to solve the composite problems as an iterative technique.

A simplified fuzzy logic rule has been provided by [40] to improve the GMDH neural network:

$$\text{If } x_1 = F_{k1} \text{ and } x_2 = F_{k2}, \text{ then output } y = w_k$$

2.6. Gray Wolf Optimization

The GWO algorithm, which is a new meta-heuristic algorithm based on swarm intelligence evolutionary, is proposed by Mirjalili et al. [42].

The GWO is inspired by grey wolves. The four types of grey wolves are hired as alpha, beta, delta, and omega to replicate the hierarchy of management. On the other side, the notable steps of grey wolves (encircling prey, hunting, attacking prey, and searching for prey) are performed during the operation [43].

Encircling prey: The encircling behaviour of each agent of the group is computed by the following mathematical formula:

$$\vec{d} = \left| c \cdot \vec{x}_p - \vec{x} \right| \quad (7)$$

$$\vec{x}^{\rightarrow t+1} = \vec{x}_p - \vec{a} \cdot \vec{d} \quad (8)$$

The vectors \vec{a} and \vec{c} are computed as the following formula:

$$\begin{cases} \vec{a} = 2l \cdot r_1 \\ \vec{c} = 2 \cdot r_2 \end{cases} \quad (9)$$

Hunting: For a mathematical simulation of the hunting behaviour of grey wolves, it is assumed that α , β , and δ have better information about the possible location of the prey.

$$\begin{cases} d_\alpha = \left| \vec{c}_1 \cdot \vec{x}_\alpha - \vec{x} \right| \\ d_\beta = \left| \vec{c}_2 \cdot \vec{x}_\beta - \vec{x} \right| \\ d_\delta = \left| \vec{c}_3 \cdot \vec{x}_\delta - \vec{x} \right| \end{cases}, \quad \begin{cases} \vec{x}_1 = \vec{x}_\alpha - \vec{a}_1 \cdot (\vec{d}_\alpha) \\ \vec{x}_2 = \vec{x}_\beta - \vec{a}_2 \cdot (\vec{d}_\beta) \\ \vec{x}_3 = \vec{x}_\delta - \vec{a}_3 \cdot (\vec{d}_\delta) \end{cases} \quad (10)$$

$$\vec{x}_{(t+1)} = \frac{\vec{x}_1 + \vec{x}_2 + \vec{x}_3}{3} \quad (11)$$

Attacking the prey and searching for the prey: The \vec{a} is a random value in the interval $[-2a, 2a]$. When the random value is less than 1, the grey wolves are enforced to attack the prey and if the random value is greater than 1, the grey wolves are forced to diverge from the prey.

2.7. Error Indicators

In order to assess the accuracy and reliability of the proposed forecasting model, different error indicators have been used in this paper: The mean absolute percentage error (MAPE), the sum squared error (SSE), the root mean squared error (RMSE), and the mean absolute error (MAE). All error indicators are based on error percentage (unitless).

$$MAPE = \frac{100}{m} \sum_{i=1}^m \left| \frac{x_{real_i} - x_{for_i}}{x_{real_i}} \right| \quad (12)$$

$$SSE = \sum_{i=1}^m (x_{real_i} - x_{for_i})^2 \quad (13)$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (x_{real_i} - x_{for_i})^2} \quad (14)$$

$$MAE = \frac{1}{m} \sum_{i=1}^m |x_{real_i} - x_{for_i}| \quad (15)$$

where x_{real_i} and x_{for_i} are the actual value and predicted value, respectively. m is the number of data.

3. Results and Discussion

As discussed in the previous sections, in order to evaluate the performance and efficiency of the proposed method, the real on-shore SCADA dataset for two wind turbines (WT1 and WT2) has been exploited in this paper. Regarding the framework of the proposed model (Figure 3), several frequencies of wind speed ($IMF1^{Speed}$, $IMF2^{Speed}$, $IMF3^{Speed}$, $IMF4^{Speed}$, Res^{Speed}) and wind turbine power ($IMF1^{Power}$, $IMF2^{Power}$, $IMF3^{Power}$, and $IMF4^{Power}$, and Res^{Power}) are considered as input parameters of the model:

$$\text{Original and decomposition signals} \left\{ \begin{array}{l} WP_{(t-1)}, WP_{(t-2)}, WP_{(t-3)}, WP_{(t-4)}, WP_{(t-5)} \\ WS_{(t-1)}, WS_{(t-2)}, WS_{(t-3)}, WS_{(t-4)}, WS_{(t-5)} \\ IMF1^{Power}_{(t-1)}, IMF2^{Power}_{(t-1)}, IMF3^{Power}_{(t-1)}, IMF4^{Power}_{(t-1)}, Res^{Power}_{(t-1)}, \\ IMF1^{Speed}_{(t-1)}, IMF2^{Speed}_{(t-1)}, IMF3^{Speed}_{(t-1)}, IMF4^{Speed}_{(t-1)}, Res^{Speed}_{(t-1)} \end{array} \right.$$

In addition, several combined forecasting models such as MI-CNN [44], MRMR-HNES [45], MI-CNEA [46], GRNN, and FGMDH have been applied to measure the performance of the proposed model. First, the 1-year-dataset is selected as the prediction model data and the predicted results are calculated with the error indicators presented in the previous section. The results are shown in Table 1.

Table 1. Comparison of the wind turbine power forecasting errors of the models for two wind turbines (WT1 and WT2).

Forecasting Models	Time Step	Error Criteria							
		RMSE		SSE		MAE		MAPE	
		WT1	WT2	WT1	WT2	WT1	WT2	WT1	WT2
GRNN	10-min	0.315	0.29	125.011	102.29	0.299	0.265	18.516	20.818
	1-h	0.381	0.387	133.366	124.971	0.312	0.365	21.815	22.716
FGMDH	10-min	0.236	0.241	18.533	19.084	0.217	0.228	9.232	11.883
	1-h	0.288	0.272	64.265	54.079	0.276	0.231	10.237	12.11
MI-CNN [44]	10-min	0.056	0.072	13.35	12.816	0.045	0.033	4.262	3.981
	1-h	0.094	0.085	40.58	41.732	0.068	0.038	4.837	5.002
MRMR-HNES [45]	10-min	0.149	0.184	11.922	9.818	0.143	0.166	4.523	5.105
	1-h	0.167	0.202	44.514	51.165	0.149	0.171	5.592	5.235
MI-CNEA [46]	10-min	0.225	0.219	14.368	13.127	0.215	0.187	6.815	5.754
	1-h	0.253	0.226	53.122	56.464	0.224	0.203	6.94	7.402
Proposed Model	10-min	0.013	0.013	10.706	8.069	0.012	0.01	2.856	3.012
	1-h	0.026	0.024	36.594	34.079	0.032	0.012	3.208	3.516

In Table 1, the prediction results are calculated for two wind turbines in two different time steps (10-min and 1-h). The wind turbine power production is highly dependent on the wind speed. On the other hand, the wind speeds vary greatly on different days thus, the forecasting time intervals for power production have been chosen amongst the days of the four following months: February, May, August, and November which have the highest fluctuation values for power production.

According to the results in Table 1, the performance of the proposed model is better than the other provided models in different time steps and wind turbines. In addition, the results indicate that the performance of the forecasting models in the 10-min time step is better than the 1-h time step. Table 2 and Figure 7 indicate the results of the proposed forecasting model and other models for wind turbine power production forecasting (WT1).

Table 2. Comparison of the wind turbine power forecasting errors of the models for the different seasons of a year.

Different Seasons	Error Criteria	GRNN		FGMDH		MI CNN [44]		MRMR HNES [45]		MI CNEA [46]		Proposed Model	
		1-h	10-min	1-h	10-min	1-h	10-min	1-h	10-min	1-h	10-min	1-h	10-min
Winter	RMSE	0.482	0.414	0.326	0.318	0.122	0.083	0.173	0.153	0.262	0.226	0.015	0.013
	SSE	171.315	46.555	37.336	16.265	19.820	9.352	28.819	11.474	22.872	13.077	17.634	6.317
	MAE	0.445	0.396	0.306	0.277	0.079	0.049	0.147	0.144	0.229	0.224	0.013	0.012
	MAPE	22.458	23.572	12.536	9.844	5.315	4.747	6.939	4.478	10.392	6.748	3.937	2.504
Spring	RMSE	0.383	0.401	0.309	0.298	0.110	0.086	0.163	0.147	0.244	0.232	0.014	0.012
	SSE	131.719	51.513	40.546	15.327	19.608	7.342	33.549	11.493	28.090	14.120	11.519	5.306
	MAE	0.363	0.359	0.284	0.275	0.074	0.054	0.147	0.138	0.222	0.208	0.019	0.012
	MAPE	20.445	19.819	15.049	11.818	5.369	4.953	7.258	4.746	10.849	7.153	4.125	2.263
Summer	RMSE	0.341	0.313	0.302	0.291	0.081	0.064	0.169	0.143	0.253	0.246	0.014	0.012
	SSE	121.481	37.914	41.151	13.011	12.780	5.311	13.907	7.453	28.994	12.030	11.134	3.006
	MAE	0.313	0.264	0.274	0.288	0.055	0.040	0.148	0.143	0.223	0.216	0.017	0.012
	MAPE	21.516	16.607	15.479	13.454	6.349	4.348	8.583	4.944	9.810	7.450	4.019	3.917
Fall	RMSE	0.372	0.328	0.285	0.268	0.118	0.074	0.163	0.162	0.247	0.244	0.018	0.014
	SSE	91.819	34.725	53.183	26.130	22.378	9.427	24.238	10.587	37.390	18.337	13.238	6.263
	MAE	0.415	0.307	0.239	0.234	0.080	0.040	0.161	0.137	0.243	0.207	0.015	0.014
	MAPE	24.160	18.504	18.710	11.366	7.419	3.498	8.141	4.119	10.487	6.214	3.015	2.903

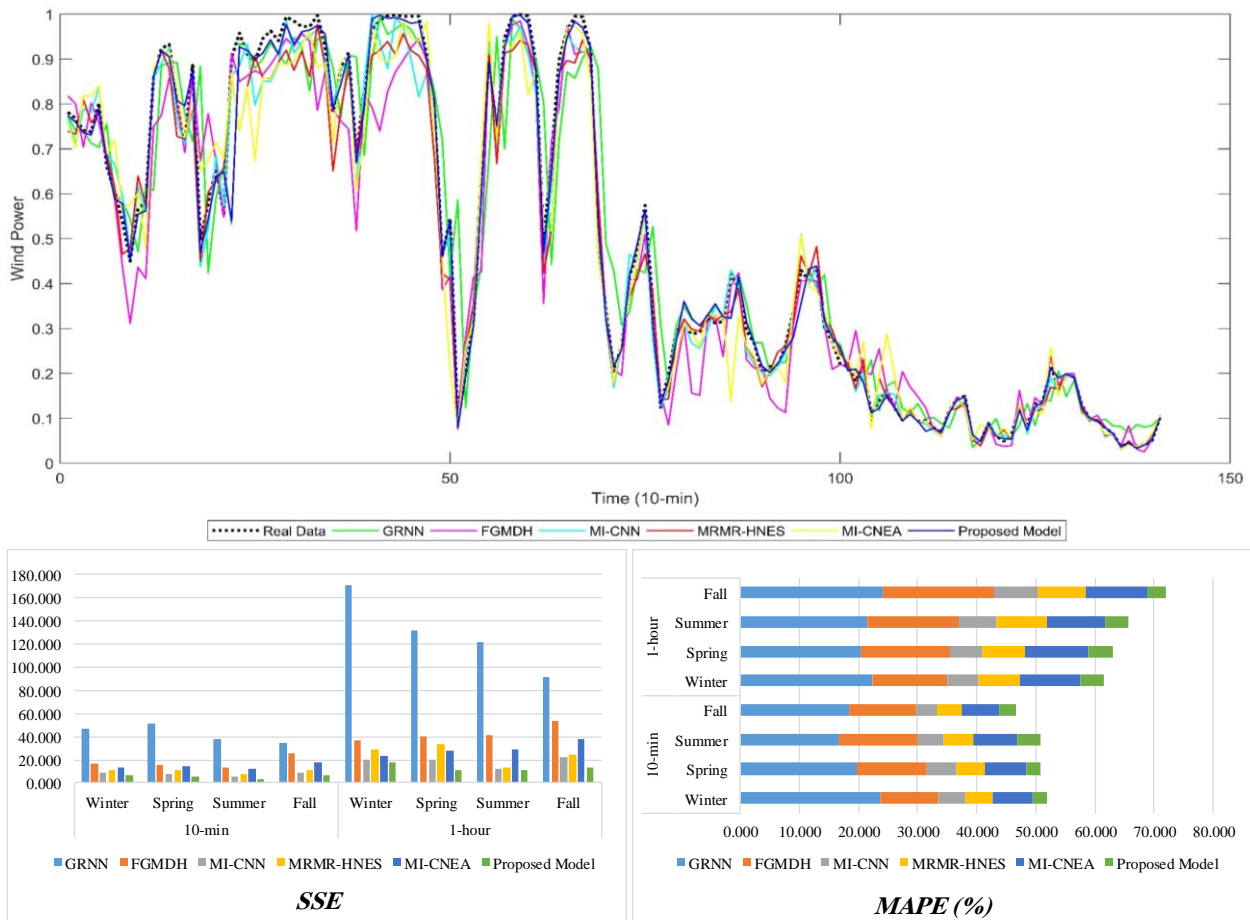


Figure 7. The wind turbine power forecasting results of the comparative models for WT1.

According to the results of Table 2 and Figure 7, the performance of the forecasting models has been evaluated in different time steps (10-min and 1-h) and seasons (winter, spring, summer, and fall). Based on these results, the proposed model can predict the wind turbine power more reliably and highly accurately multiple times ahead compared to the other valid forecasting models (GRNN, FGMDH, MI-CNN, MRMR-HNES, and MI-CNEA).

4. Conclusions

Considering the highly volatile and nonlinear process of wind turbines power production, a hybrid intelligent system to improve the accuracy and efficiency of wind turbine power prediction has been proposed. For the initial step, the hybrid K-means-LOF and EMD methods have been applied as a pre-processing step for removing the outliers and decomposition of the SCADA data, respectively. Then, the processed data was given to the forecasting model (FGMDH) and the future power of the wind turbine has been calculated. Furthermore, in order to complete the proposed model as a parallel calculation, the GWO algorithm has been used as an optimization method to optimize the FGMDH parameters. In this study, the SCADA data for two wind turbines in the real wind farm located in Sweden has been used to measure the performance and reliability of the proposed model.

The new forecasting model has been applied to predict the power of wind turbines for two time-intervals ahead (10-min and 1-h) in 1 year and different seasons. The obtained results pinpointed that the performance of the proposed method (EMD-FGMDH-GWO) at different time intervals has a high accuracy and reliability than many other available methods such as GRNN, FGMDH, MI-CNN, MRMR-HNES, and MI-CNEA. The MAPE error indicator obtained for GRNN, FGMDH, MI-CNN, MRMR-HNES, MI-CNEA and the proposed model is equal to 20.818, 11.883, 3.981, 5.105, 5.754, and 3.012, respectively. In

addition, the proposed forecasting model can be extended and applied for different energy sources to maximize the use of renewable energy sources and better manage their use. In future studies, the proposed model will be extended to analyze different energy sources in one area simultaneously.

Author Contributions: Conceptualization, A.H.; data curation, A.H., M.M.N., and M.N.; investigation, A.H., M.M.N., and M.N.; methodology, A.H., M.N., and M.M.N.; validation, A.H., M.M.N., M.N., D.A.G., and F.K.; visualization, A.H., M.M.N., D.A.G., and F.K.; writing—original draft, A.H., D.A.G., F.K., and L.B.T.; writing—review and editing, A.H., D.A.G., F.K., and L.D.S.; supervision, D.A.G., F.K., L.D.S., and L.B.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Okumus, I.; Dinler, A. Current status of wind energy forecasting and a hybrid method for hourly predictions. *Energy Convers. Manag.* **2016**, *123*, 362–371. [[CrossRef](#)]
- Zhao, Y.; Ye, L.; Li, Z.; Song, X.; Lang, Y.; Su, J. A novel bidirectional mechanism based on time series model for wind power forecasting. *Appl. Energy* **2016**, *177*, 793–803. [[CrossRef](#)]
- Karki, R.; Hu, P.; Billinton, R. A simplified wind power generation model for reliability evaluation. *IEEE Trans. Energy Convers.* **2006**, *21*, 533–540. [[CrossRef](#)]
- Li, S.; Wunsch, D.C.; O’Hair, E.A.; Giesselmann, M.G. Using neural networks to estimate wind turbine power generation. *IEEE Trans. Energy Convers.* **2001**, *16*, 276–282. [[CrossRef](#)]
- Xia, S.; Chan, K.W.; Luo, X.; Bu, S.; Ding, Z.; Zhou, B. Optimal sizing of energy storage system and its cost-benefit analysis for power grid planning with intermittent wind generation. *Renew. Energy* **2018**, *122*, 472–486. [[CrossRef](#)]
- Akçay, H.; Filik, T. Short-term wind speed forecasting by spectral analysis from long-term observations with missing values. *Appl. Energy* **2017**, *191*, 653–662. [[CrossRef](#)]
- Toh, G.K.; Gooi, H.B. Incorporating forecast uncertainties into EENS for wind turbine studies. *Electr. Power Syst. Res.* **2011**, *81*, 430–439. [[CrossRef](#)]
- Ortega-Vazquez, M.A.; Kirschen, D.S. Estimating the spinning reserve requirements in systems with significant wind power generation penetration. *IEEE Trans. Power Syst.* **2009**, *24*, 114–124. [[CrossRef](#)]
- Wu, J.; Zhang, B.; Li, H.; Li, Z.; Chen, Y.; Miao, X. Statistical distribution for wind power forecast error and its application to determine optimal size of energy storage system. *Int. J. Electr. Power Energy Syst.* **2014**, *55*, 100–107. [[CrossRef](#)]
- Usaola, J. Probabilistic load flow with correlated wind power injections. *Electr. Power Syst. Res.* **2010**, *80*, 528–536. [[CrossRef](#)]
- Balasubramaniam, K.; Saraf, P.; Hadidi, R.; Makram, E.B. Energy management system for enhanced resiliency of microgrids during islanded operation. *Electr. Power Syst. Res.* **2016**, *137*, 133–141. [[CrossRef](#)]
- Bruninx, K.; Delarue, E. A statistical description of the error on wind power forecasts for probabilistic reserve sizing. *IEEE Trans. Sustain. Energy* **2014**, *5*, 995–1002. [[CrossRef](#)]
- Zhang, Z.S.; Sun, Y.Z.; Gao, D.W.; Lin, J.; Cheng, L. A versatile probability distribution model for wind power forecast errors and its application in economic dispatch. *IEEE Trans. Power Syst.* **2013**, *28*, 3114–3125. [[CrossRef](#)]
- Tascikaraoglu, A.; Sanandaji, B.M.; Chicco, G.; Cocina, V.; Spertino, F.; Erdinc, O.; Paterakis, N.G.; Catalao, J.P. A short-term spatio-temporal approach for Photovoltaic power forecasting. In Proceedings of the 2016 Power Systems Computation Conference (PSCC), Genoa, Italy, 20–24 June 2016. [[CrossRef](#)]
- Li, S.; Dong, W.; Huang, J.; Wu, Z.; Zhang, H. Wind power system reliability sensitivity analysis by considering forecast error based on non-standard third-order polynomial normal transformation method. *Electr. Power Syst. Res.* **2019**, *167*, 122–129. [[CrossRef](#)]
- Heydari, A.; Astiaso Garcia, D.; Keynia, F.; Bisegna, F.; De Santoli, L. A novel composite neural network based method for wind and solar power forecasting in microgrids. *Appl. Energy* **2019**, *251*, 113353. [[CrossRef](#)]
- Cui, Y.; Bangalore, P.; Bertling Tjernberg, L. An Anomaly Detection Approach Based on Machine Learning and SCADA Data for Condition Monitoring of Wind Turbines. In Proceedings of the 2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), Boise, ID, USA, 24–28 June 2018; pp. 1–6.
- Heydari, A.; Keynia, F. A new intelligent heuristic combined method for short-term electricity price forecasting in deregulated markets. *Aust. J. Electr. Electron. Eng.* **2016**, *13*, 258–267. [[CrossRef](#)]
- Zhang, X.; Wang, J.; Zhang, K. Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by Cuckoo search algorithm. *Electr. Power Syst. Res.* **2017**, *146*, 270–285. [[CrossRef](#)]

20. Kouhi, S.; Keynia, F.; Najafi Ravadanegh, S. A new short-term load forecast method based on neuro-evolutionary algorithm and chaotic feature selection. *Int. J. Electr. Power Energy Syst.* **2014**, *62*, 862–867. [[CrossRef](#)]
21. Amjady, N.; Keynia, F.; Zareipour, H. Short-term wind power forecasting using ridgelet neural network. *Electr. Power Syst. Res.* **2011**, *81*, 2099–2107. [[CrossRef](#)]
22. Han, Q.; Meng, F.; Hu, T.; Chu, F. Non-parametric hybrid models for wind speed forecasting. *Energy Convers. Manag.* **2017**, *148*, 554–568. [[CrossRef](#)]
23. Pelajo, J.C.; Brandão, L.E.T.; Gomes, L.L.; Klotzle, M.C. Wind farm generation forecast and optimal maintenance schedule model. *Wind Energy* **2019**, *22*, 1872–1890. [[CrossRef](#)]
24. Osório, G.J.O.; Matias, J.C.O.; Catalão, J.P.S. Hybrid evolutionary-adaptive approach to predict electricity prices and wind power in the short-term. In Proceedings of the 2014 Power Systems Computation Conference, Wroclaw, Poland, 18–22 August 2014. [[CrossRef](#)]
25. Gallego-Castillo, C.; Cuerva-Tejero, A.; Bessa, R.J.; Cavalcante, L. Wind power probabilistic forecast in the reproducing kernel Hilbert space. In Proceedings of the 2016 Power Systems Computation Conference (PSCC), Genoa, Italy, 11 August 2016; pp. 1–7. [[CrossRef](#)]
26. Xiao, L.; Shao, W.; Yu, M.; Ma, J.; Jin, C. Research and application of a hybrid wavelet neural network model with the improved cuckoo search algorithm for electrical power system forecasting. *Appl. Energy* **2017**, *198*, 203–222. [[CrossRef](#)]
27. Shi, K.; Qiao, Y.; Zhao, W.; Wang, Q.; Liu, M.; Lu, Z. An improved random forest model of short-term wind-power forecasting to enhance accuracy, efficiency, and robustness. *Wind Energy* **2018**, *21*, 1383–1394. [[CrossRef](#)]
28. Doan, V.Q.; Kusaka, H.; Matsueda, M.; Ikeda, R. Application of mesoscale ensemble forecast method for prediction of wind speed ramps. *Wind Energy* **2019**, *22*, 499–508. [[CrossRef](#)]
29. Duan, J.; Wang, P.; Ma, W.; Tian, X.; Fang, S.; Cheng, Y. Short-term wind power forecasting using the hybrid model of improved variational mode decomposition and Correntropy Long Short-term memory neural network. *Energy* **2021**, *214*, 118980. [[CrossRef](#)]
30. Yildiz, C.; Acikgoz, H.; Korkmaz, D.; Budak, U. An improved residual-based convolutional neural network for very short-term wind power forecasting. *Energy Convers. Manag.* **2021**, *228*, 113731. [[CrossRef](#)]
31. Jafarzadeh, S.; Manjili, S.; Mardani, A.; Kamali, M. An extended new approach for forecasting short-term wind power using modified fuzzy wavelet neural network: A case study in wind power plant. *Energy* **2021**, *223*, 120052. [[CrossRef](#)]
32. Groppi, D.; De Santoli, L.; Cumo, F.; Astiaso Garcia, D. A GIS-based model to assess buildings energy consumption and usable solar energy potential in urban areas. *Sustain. Cities Soc.* **2018**, *40*, 546–558. [[CrossRef](#)]
33. De Santoli, L.; Mancini, F.; Astiaso Garcia, D. A GIS-based model to assess electric energy consumptions and usable renewable energy potential in Lazio region at municipality scale. *Sustain. Cities Soc.* **2019**, *46*, 101413. [[CrossRef](#)]
34. Mancini, F.; Nastasi, B. Solar energy data analytics: PV deployment and land use. *Energies* **2020**, *13*, 417. [[CrossRef](#)]
35. Nastasi, B.; De Santoli, L.; Albo, A.; Bruschi, D.; Lo Basso, G. RES (Renewable Energy Sources) availability assessments for Eco-fuels production at local scale: Carbon avoidance costs associated to a hybrid biomass/H₂ NG-based energy scenario. *Energy Procedia* **2015**, *81*, 1069–1076. [[CrossRef](#)]
36. Huang, Q.; Cui, Y.; Bertling Tjernberg, L.; Bangalore, P. Wind turbine health assessment framework based on power analysis using machine learning method. In Proceedings of the 2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), Bucharest, Romania, 29 September–2 October 2019; pp. 1–5. [[CrossRef](#)]
37. Breunig, M.M.; Kriegel, H.P.; Ng, R.T.; Sander, J. LOF: Identifying density-based local outliers. *SIGMOD Rec. (ACM Spec. Interes. Gr. Manag. Data)* **2000**, *29*, 93–104. [[CrossRef](#)]
38. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **1998**, *454*, 903–995. [[CrossRef](#)]
39. Huang, N.E.; Wu, M.L.; Qu, W.; Long, S.R.; Shen, S.S.P. Applications of Hilbert-Huang transform to non-stationary financial time series analysis. *Appl. Stoch. Model. Bus. Ind.* **2003**, *19*, 245–268. [[CrossRef](#)]
40. Ohtani, T.; Ichihashi, H.; Miyoshi, T.; Nagasaka, K. Orthogonal and successive projection methods for the learning of neurofuzzy GMDH. *Inf. Sci.* **1998**, *110*, 5–24. [[CrossRef](#)]
41. Najafzadeh, M. Neurofuzzy-based GMDH-PSO to predict maximum scour depth at equilibrium at culvert outlets. *J. Pipeline Syst. Eng. Pr.* **2016**, *7*, 06015001. [[CrossRef](#)]
42. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey Wolf Optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [[CrossRef](#)]
43. Song, J.; Wang, J.; Lu, H. A novel combined model based on advanced optimization algorithm for short-term wind speed forecasting. *Appl. Energy* **2018**, *215*, 643–658. [[CrossRef](#)]
44. Keynia, F. A new feature selection algorithm and composite neural network for electricity price forecasting. *Eng. Appl. Artif. Intell.* **2012**, *25*, 1687–1697. [[CrossRef](#)]
45. Amjady, N.; Keynia, F. Application of a new hybrid neuro-evolutionary system for day-ahead price forecasting of electricity markets. *Appl. Soft Comput.* **2010**, *10*, 784–792. [[CrossRef](#)]
46. Amjady, N.; Keynia, F. Day ahead price forecasting of electricity markets by a mixed data model and hybrid forecast method. *Int. J. Electr. Power Energy Syst.* **2008**, *30*, 533–546. [[CrossRef](#)]