Using Data-Mining Approaches for Wind Turbine Power Curve Monitoring: A Comparative Study

Meik Schlechtingen, Ilmar Ferreira Santos, and Sofiane Achiche

***Abstract*—Four data-mining approaches for wind turbine power curve monitoring are compared. Power curve monitoring can be applied to evaluate the turbine power output and detect deviations, causing financial loss. In this research, cluster center fuzzy logic, neural network, and k-nearest neighbor models are built and their performance compared against literature. Recently developed adaptive neuro-fuzzy-interference system models are set up and their performance compared with the other models, using the same data. Literature models often neglect the influence of the ambient temperature and the wind direction. The ambient temperature can influence the power output up to 20%. Nearby obstacles can lower the power output for certain wind directions. The approaches proposed in literature and the ANFIS models are compared by using wind speed only and two additional inputs. The comparison is based on the mean absolute error, root mean squared error, mean absolute percentage error, and standard deviation using data coming from three pitch regulated turbines rating 2 MW each. The ability to highlight performance deviations is investigated by use of real measurements. The comparison shows the decrease of error rates and of the ANFIS models when taking into account the two additional inputs and the ability to detect faults earlier.**

***Index Terms*—Condition monitoring, data mining, fuzzy neural networks, machine learning, neural networks, power generation, power system faults, signal analysis, wind energy.**

NOMENCLATURE

ANFIS	Adaptive neuro-fuzzy interference system.
CCFL	Cluster center fuzzy logic.
k-NN	k-nearest neighbor.
M5P	Quinlan' M5 algorithm for including trees.
MAE	Mean absolute error.
MAPE	Mean absolute percentage error.
MF	Membership function.
MLP	Multilayer-perceptron.
NN	Neural network.
NR	Normal range.
REP	Representative.
RMS	Root mean squared error.
SCADA	Supervisory control and data acquisition.
SD	Standard deviation.
WEC	Wind energy converter.

Manuscript received October 02, 2012; revised December 06, 2012; accepted January 05, 2013. Date of publication February 14, 2013; date of current version June 17, 2013.

M. Schlechtingen is with the Department of Technical Operation Wind Offshore, EnBW Erneuerbare Energien GmbH, 20459 Hamburg, Germany (e-mail: m.schlechtingen@enbw.com).

I. F. Santos is with the Department of Mechanical Engineering, Section of Solid Mechanics, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark.

S. Achiche is with the Department of Mechanical Engineering, Machines Design Section, Ecole Polytechnique de Montréal, Montréal, QC, H3C 3A7, Canada.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSTE.2013.2241797

I. INTRODUCTION

IN THE past decades, the cumulated worldwide installed capacity of wind energy converters (WECs) grew exponentially. One of the reasons is the achieved reduction in cost of energy that has today reached a level where it is almost comparable to conventionally generated power from coal and gas fired power plants. More and more wind turbine operators decide to trade their energy directly on the electricity market. For this purpose and to keep the cost of energy down and increase profit margins, operators need to be able to prognosticate the performance of their turbines more accurately. In case of decreased turbine performance, operators may be unable to deliver their traded amount of energy and consequentially have to pay fines. Furthermore, financial loss is generated as the power output of the turbine is lower than expected and the revenue is hence missing on the balance sheet. Here, power curve monitoring can serve as an effective method to evaluate the performance as power curves for WECs describe the essential relation between wind speed and electrical power output [1]. Detected decrease allows the operator to take action to identify the root cause and improve performance.

Different models were proposed in the past to estimate wind turbine power curves for performance evaluation. The basic idea of all model approaches in this context is to identify closely related signals (e.g., the wind speed) to use them to build a model of the power output. After model training (learning the model the input–output, e.g., wind speed—power output relation), the model is kept fixed and applied in the following using the inputs to obtain an expectation of the output. The prediction error can then be an indicator for anomaly—the prediction error is defined here as the difference between the model's output (expectation) and the real measurement.

In 1997, Li *et al.* [2] presented a method using multilayer perceptron (MLP) neural networks (NNs) to predict wind power generation of stall regulated wind turbines. NNs can learn nonlinear relationships between input and output data sets by use of activation functions within the hidden neurons. However, it uses a black box approach to globally fit a single function to the data and thereby losing insight into the problem [3].

Four years later the same authors published a comparative analysis of regression and artificial NN models for wind turbine power curve estimation [4]. The comparison showed that NN models perform better than regression models. In 2008 Üstüntaş and Duran Şahin applied cluster center fuzzy logic (CCFL) modeling for power curve estimation [5]. In this method, the number of cluster centers is determined and their initial values for initializing iterative optimization are based on clustering algorithms. [5]. The CCFL model is compared to a polynomial model estimated with the least squares algorithm. It is concluded that the root mean squared (RMS) error is much lower when using the fuzzy logic methodology [5]. However, the clustering approach puts limits on the type of MF used and thus the accuracy of the model depends mainly on the number of cluster centers. A probabilistic (Copula) model of wind turbine performance was proposed by Stephen *et al.* [6] in 2011, but the performance was not quantified in terms of the metrics used here.

It is worth noting that the modeling methodologies presented in [1]–[6] were applied to stall regulated (fixed pitch angle) turbines only. The pitch angle describes the angle of the rotor blade, which adjustment can be used to control the aerodynamic lift. Today, turbines are pitch regulated (variable pitch angle). Stall regulated wind turbines have decreased power output above rated wind speed (stall region), while pitch regulated wind turbines maintain constant power in this region up to the maximum operating wind speed (cut-out wind speed).

In 2009, Kusiak, *et al.* compared five nonparametric models for monitoring wind farm power and compared their performances [7]. These were: MLP NNs, Quinlan's M5 algorithm for including trees (M5P), representative (REP) tree, bagging tree, and k-nearest neighbor (k-NN) models. Model trees are a type of decision tree with linear regression functions at the leaves, and form the basis of a recent successful technique for predicting continuous numeric values. They can be applied to classification problems by employing a standard method of transforming a classification problem into a problem of function approximation [8]. In the k-NN algorithm, each new instance is compared with existing ones using a distance metric and the closest existing instances are used to assign the class to the new one [9].

The models were applied, tested, and evaluated at a wind power plant consisting of pitch regulated wind turbines with the aim of abnormal power output detection. The k-NN models performed best. The authors also compared the performance of the models to least squares and maximum likelihood parametric models. The least squares parametric model showed good performances [10], but the performance of those models is not quantified in terms of the mean absolute error (MAE) or the RMS.

Among the several comparative studies found in literature, the best performing and recommended models for wind turbine power curve monitoring are:

1) CCFL models [5].
2) NN models [4], [7], [10].
3) k-NN models [7], [10].

This paper presents the findings of a power curve modeling performance study using data from modern pitch regulated wind turbines. For this purpose, CCFL, NN, and k-NN models are set up in accordance with [4], [5], [7], and [10], and their performance is compared to the performance studies reported in these references.

In addition, adaptive neuro-fuzzy interference system (ANFIS) [11] models are developed and their performance evaluated and compared. Earlier computational studies by Kusiak and Wei [12] have proven that the model for methane production prediction built by the ANFIS algorithm offered the best performance [12]. ANFIS models can learn nonlinear signal relations by setting up a set of fuzzy rules and tuning the membership function (MF) parameters in a training phase. In [11], it is shown that in comparison to NN models, fewer parameters must be trained in ANFIS models, leading generally to faster training. Furthermore, the output of NNs is difficult to backtrack (black box model) and *a priori* knowledge about the system is difficult to be incorporated. Here ANFIS models have a major advantage, due to their output being based on linguistic rules and tuneable MFs.

Modeling approaches proposed in literature tend to neglect the influence of the ambient temperature and the wind direction as input parameters. In Section II, it is shown that their influence is of nonnegligible order and, therefore, including them is expected to decrease the prediction error. The performance of the initially developed models is, therefore, also compared to models considering the two additional inputs.

In this framework, the main original contribution of this paper lies in the direct comparison of different model approaches found in the literature based on data sets reflecting modern turbine behavior (pitch regulated turbines) as well as a comparison to newly developed ANFIS models. Moreover, the models are enhanced by considering the ambient temperature and the wind direction to further improve the prediction performance. Performance is measured in terms of the MAE and RMS. Furthermore, the performance of the extended models is additionally evaluated using the mean absolute percentage error (MAPE) and the standard deviation (SD) of the prediction error. A further performance measure used is the ability to highlight abnormal turbine performance. For this scope, the trained models are applied to a data set containing real measured performance deviations.

In Section II the general power curve properties are discussed and the influencing factors identified based on theory. Moreover, it is mentioned where deviations from theory may stem from. In Section III the data sets and the data preprocesses are described, used to filter invalid data and nonoperational periods as well as remove signal lags. Section IV introduces the different model types and training algorithms. Optimal settings for the different models (e.g., number of neurons or cluster centers) are selected and discussed. Section V summarizes the performance comparison based on the statistical metrics using literature and our own model results. Finally, conclusions are drawn in Section VI.

II. General Power Curve Properties

In this section, the general power curve properties are discussed in order to identify the most influential factors on the power output.

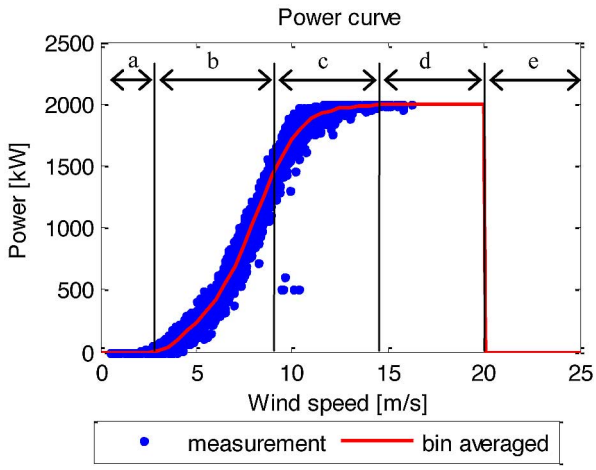

Fig. 1. Power curve properties.

The theoretical power output of a wind turbine can be expressed by the following equation:

$$P = 0.5\rho A C_p V^3 \tag{1}$$

where P is the power output, C_p the power coefficient, ρ the air density, V the wind speed, and A the swept rotor area. The variable having the largest effect on the power output is the wind speed as it is powered by three.

The power coefficient C_p is a function of the blade pitch angle which is controlled by the turbine controller to regulate the power output. Large pitch angles lead to lower power coefficients and thus lower power output [13].

The air density ρ is defined as $\rho = p/(R \cdot T)$ with p being the absolute air pressure, R the specific gas constant, and T the absolute air temperature. "p" is a function of the altitude of the wind turbine location and the fluctuations due to weather phenomena. The pressure is constant when considered in relation to the altitude at the turbine location whereas it varies by about 10% as a function of weather phenomena. The specific gas constant is as function of the humidity. Wet air has a specific gas constant that is about 2% larger than dry air. The air temperature may vary in the range of -20 °C to 30 °C (20%). Hence the temperature has the largest influence on the air density and consequentially also influences the power output. This effect is often neglected in the literature when developing models for the power curve estimation, leading to larger prediction errors. Furthermore, the wind direction may have an influence due to terrain and shading effects of neighboring turbines or nearby obstacles.

A typical power curve of a pitch regulated wind turbine is shown in Fig. 1.

As illustrated in Fig. 1, the power curve can be divided into five regions:

a) wind speed $<$ cut-in speed—the cut-in speed is the wind speed value at which the turbine starts operating; power output zero or negative; pitch angle about 90°
b) wind speed $>$ cut-in speed $<$ rated speed; power output according to (1); pitch angle between$0 \ldots -5°$
c) wind speed around rated speed; power output according to (1); pitch angle between $-5 \ldots 5°$
d) wind speed $>$ rated speed; power output according to (1); pitch angle between $5 \ldots 25°$
e) wind speed $>$ cut-out speed; power output zero or negative; pitch angle about 90°

In general, the power curve measured can deviate from the theoretical value due to:

- Inaccuracies in wind speed measurement; for example, the point of measurement can be situated on top of the nacelle, whereas the power output depends on the wind speed across the rotor area.
- Air density fluctuations.
- Yaw and pitch misalignments.
- Issues related to period averaging, as the power output has a cubic dependency of the wind speed; strong wind speed fluctuations in the averaging period influences more the power output.
- Power output restrictions manually set in the controller.
- Shading effects from neighboring turbines or nearby obstacles.

Among others these influences cause the variance in the measured power curve visible in Fig. 1. Influences caused by the measurement procedure and the used equipment are addressed in an IEC standard (61400-12-1) [14]. However, in practice these recommendations are seldom met, and mostly used for initial power performance measurements performed by turbine manufacturers [6]. Models which only consider the wind speed as input will not be able to model any of the variance and are thus expected to have a larger prediction error. In this research, the air density influences as well as the shading effect are considered in order to reduce the prediction error.

III. Data Sets and Data Preprocessing

The data sets, used in this paper, come from three wind turbines of the 2-MW class and contain 149 supervisory control and data acquisition (SCADA) signals, including the power output, wind speed, wind direction, and the ambient temperature.

Two hundred twenty-nine days of data are selected for model training and validation (32 976 10-min average values in total). The data are collected from December 2010 until August 2011 and thus cover both a summer and a winter period.

The data sets are randomly divided into training and validation data sets with a ratio of 60%/40%. The validation data sets are excluded from model training and are only used for performance evaluation in terms of the MAE, RMS, MAPE, and SD.

Before training the models, the data are preprocessed according to the methodology proposed in [15] which encompasses the three following steps: 1) validity check, 2) data range check, and 3) missing data processing. Moreover, only the data where the turbine is operating (power output $>$ 25 kW) is used. The data set sizes after preprocessing are given in Table I.

The power curves of the three wind turbines considered are illustrated in Fig. 2 together with the remaining data after preprocessing.

In addition to these, a time series from a later period exists which contains an anomaly in terms of a reduction in turbine

TABLE I
REMAINING DATA SIZES AFTER DATA PREPROCESSING

Turbine	Training [10 min avg.]	Validation [10 min avg.]
WEC A	14430	9621
WEC B	13535	9024
WEC C	14327	9552

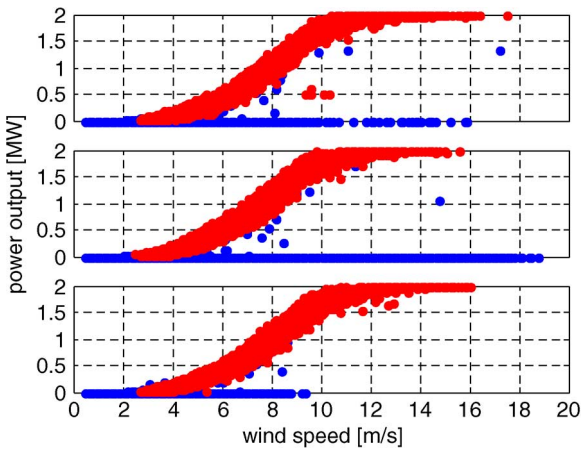

Fig. 2. Power curves derived from preprocessed data (red) and raw data (blue). Top: WEC A; middle: WEC B; bottom: WEC C.

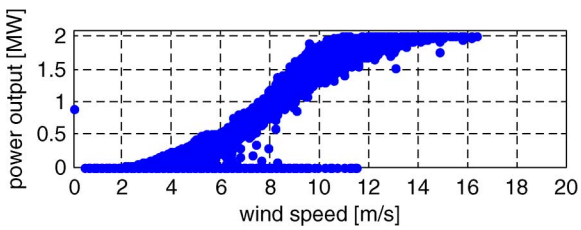

Fig. 3. Power curve during abnormal turbine performance (WEC A).

power output, caused by a wrong controller setting after a controller software update (Fig. 3).

In comparison to Fig. 2, the power curve in Fig. 3 is broader especially for wind speeds above 9 m/s. Without the comparison to the normal power curve, this type of deviation is difficult to identify by a visual check.

IV. MODELS

For each approach, two model types are set up. Type one is the one commonly used in literature (e.g., in [4], [5], [7], and [10]), whereas with type two, the ambient temperature and the wind direction are also considered. The two types are illustrated Fig. 4.

In order to evaluate the contribution of each of the additional inputs of type 2, an exhaustive search is performed by training ANFIS models with different input combinations (using the preprocessed data of WEC A). Computational time can be saved, by knowing that the input combination 1-2-3 gives the same result as 1-3-2, since the order of inputs presented to ANFIS can

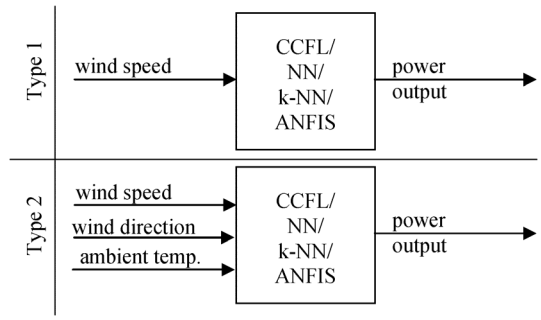

Fig. 4. Model types used for power output modeling.

TABLE II
INPUT SIGNAL IMPORTANCE IN TERMS OF THE CORRELATION COEFFICIENT

Input no.	Input name	1 Input	2 Inputs	3 Inputs
1	Wind speed	0.9939	0.9947	0.9949
2	Ambient temp			
3	Wind direction			
1	Wind speed	0.9939	0.9941	0.9949
2	Wind direction			
3	Ambient temp.			
Computational time[1)]		5.85s	14.85s	35.81

[1)] Using a standard 2.6-GHz processor

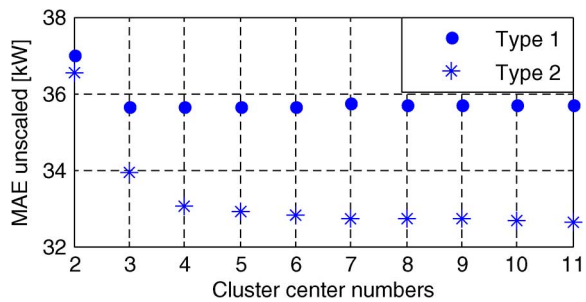

Fig. 5. Relation between the MAE and numbers of cluster centers (WEC A).

be arbitrary. As a measure of importance, the correlation coefficient of the predictions and the measurements is used. The result is summarized in Table II.

The ambient temperature contributes to the overall explained variance of the target signal with 0.0008 (0.08%), while the wind direction contributes 0.0002 (0.02%).

A. CCFL Model

The cluster center fuzzy logic (CCFL) model is set up in accordance with [5] using the algorithm proposed by [16]. Additional information can be found in [17]. The main free parameter is the number of cluster centers. In [5], eight cluster centers are used. In Fig. 5, the relation between the MAE and the number of clusters is shown for the training data set coming WEC A.

Based on Fig. 5 three cluster centers are used for model type one and six for type two. For model type two, the MAE does not decrease significantly for larger numbers of cluster centers.

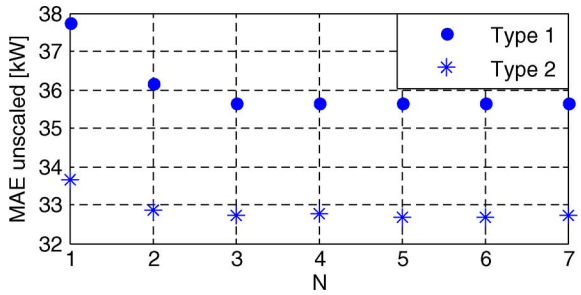

Fig. 6. Relation between the MAE and numbers of hidden neurons N (WEC A).

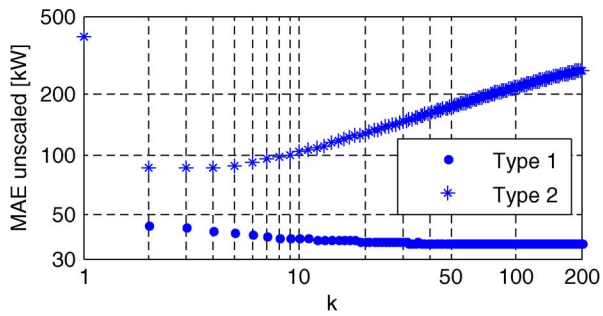

Fig. 7. Relation between the MAE and number of neighbor's k (WEC A).

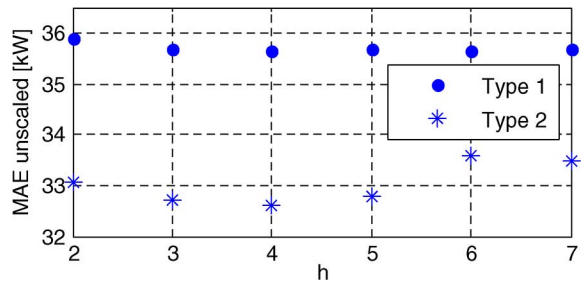

Fig. 8. Relation between the MAE and number of MF's h (WEC A).

B. NN Model

In [10], no information about the setup of the MLP NN is given.

Next to the number of hidden layers, the number of hidden neurons is an important parameter characterizing the NN structure [15]. The number of hidden layers chosen is two, as with three layers, no major performance improvement was observed, but the computational time increased drastically. In [18], it is recommended that the optimum number of neurons should be found by performing at least 10 runs where only the number of neurons is changed. The network architecture which gives the best generalization should be chosen. This procedure is also followed here. The number of hidden neurons in the hidden layer is $N \cdot n_i$, where n_i is the number of input signals and N an integer multiplier. In accordance with Fig. 6, N is chosen to be three for both model types since for N larger three the MAE does not significantly change.

The training method used is gradient decent with momentum. The advantage of this method is that not only the local gradient of the error function is calculated, but so is the general trend. Thus local minima may be survived and generalization is improved [15].

Sigmoid transfer functions are used for the hidden layer and linear transfer functions for the output layer as this proved to give better results. Next to the Sigmoid transfer functions, logarithmic and linear functions were evaluated. More information about NN models can be found in [19]–[21].

C. k-NN Model

The k-nearest neighbor model is set up according to the methodology stated in [22]. In [10], the number of neighbor's k was set to 150. Fig. 7 shows the relation between the MAE and the number of neighbors for the training data set coming from WEC A.

The number of k used for modeling of type one is 150 and three for type two. Although the performance of model type one does not change much for $k > 30$, k is chosen to be 150 because of this value is used in the reference study allowing a better comparison. More information about the k-NN method can be found in [9] and [22].

D. ANFIS Model

The adaptive neuro-fuzzy interference system model build is described in more detail, since these models are comparably young and the application to the specific problem presents a novelty. Additional information about ANFIS models as well as theory can be found in [11].

1) FIS Structure: There are two common types of fuzzy inference: the Mamdani [23] and Sugeno [24]. The main difference between the two types is in the consequent part, which can be a nonfuzzy equation (Sugeno) instead of a fuzzy linguistic value (Mamdani) [25]. In this work, the Sugeno type is used as it has fuzzy sets only involved in the premise part. However, the consequent part can be a nonfuzzy equation. Due to the qualifiers on the premise parts, each fuzzy if–then rule can be viewed as a local description of the system under consideration [11]. Moreover, the Sugeno type is known to be computationally more efficient and has guaranteed continuity of the output surface, which makes it well suited for power curve prediction.

2) Membership Functions: Membership functions can be represented by an arbitrary curve whose shape is defined as a function that suits the application. In this paper, generalized normal distribution MFs are employed for the input space and linear membership functions for the output space. The generalized normal distribution functions have the advantage of giving a broad flexibility with regards to the function shape, depending on their parameters. Furthermore, these functions assure smoothness of the transitions in the input space. The free parameters are: location, scale, and shape [25].

3) Number of Membership Functions/Rules: In a conventional fuzzy interference system (FIS), the number of rules is decided by an expert who is familiar with the system to be modeled [11]. Generally, the number of MFs can be defined for each input separately [25]. In this research, the expert for this particular task is assumed to be unavailable and the number of MFs h per input for both model types is set to three, in agreement with Fig. 8.

TABLE III
TYPE 1 MODEL PERFORMANCE COMPARISON

*	Un-scaled	Scaled to 0…100 kW			
		New models		Literature models	
	MAE	MAE	RMS	MAE	RMS
CCFL	38.72	1.94	2.78	-	*0.49*[3] [5]
NN	38.74	1.94	2.78	5.3[1] [4] 3.56[2] [7] 5.50 [10]	-
k-NN	39.03	1.95	2.80	2.15[2] [7] 3.19 [10]	-
ANFIS	38.71	1.94	2.78	-	-

[1] Rescaled from 500 kW. [2] Rescaled from total plant output. [3] Calculated from bin averaged data set. * Values given in kW

Although the MAE for type two and h equal four is smaller, the computational time drastically increases with increasing h, but the performance improvements are not significant.

4) Training Method: The training algorithm used is a hybrid learning rule as proposed in [11], consisting of a combination of gradient decent and least squares estimation. Not only can this hybrid learning approach decrease the dimension of the search space in the gradient method, but in general, it will also cut down substantially the convergence time [11].

V. PERFORMANCE COMPARISON

In this section, the performance of the models is compared using MAE, RMS, and the ability to identify abnormal turbine power output. The values coming from our own calculations are averaged among the model performance of the three turbines considered.

In order to compare the achieved performance to values stated in the literature, rescaling is required. The performance values stated in [10] and [5] are scaled to the 0. . .100-kW range. Rescaling here is done by dividing the result by the turbine rated power and multiplying by 100 kW. The performance values in this paper will be given both as scaled and nonscaled values.

A. Type 1 Model Performance Comparison

1) Performance Based on MAE and RMS: Table III summarizes the results of the performance evaluation.

The performance of all models is comparable. The chosen free parameters allow all models to map the existing relation between the power output and the wind speed. Fig. 9 shows the measured and the estimated power curve for the validation data set of WEC A using the set up ANFIS model.

The model is incapable of representing the variance of the measurements, but the general power curve shape appears to be well modeled.

The performance of the CCFL, NN, and k-NN models is better in terms of the MAE than the values reported in the literature (compare Table III). The RMS value coming from [5] is difficult to compare as here an averaged power curve is modeled. Using a similar procedure, the achieved RMS with the set up CCFL model is 0.21 kW.

In [4] and [5], data from stall regulated wind turbines are modeled. Power curves of those turbines have a broader scatter

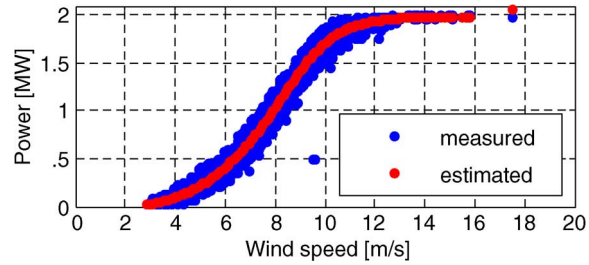

Fig. 9. Power curve of WEC A for the validation period; measured and estimated by the ANFIS model (type 1).

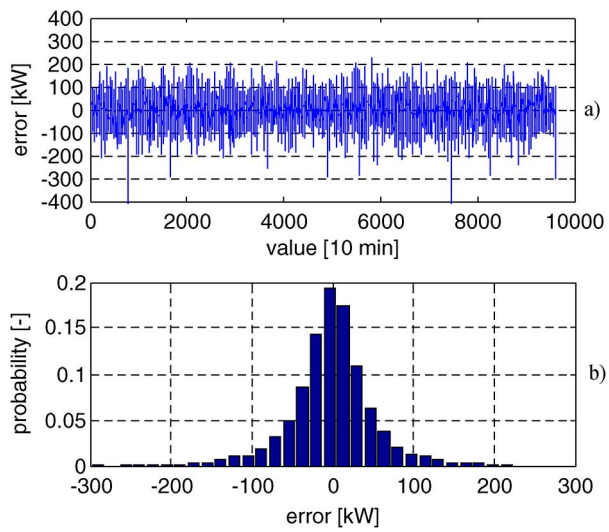

Fig. 10. (a) Prediction error over time; (b) probability distribution of the prediction error (WEC A; type 1).

above rated wind speeds, which is caused by the stall effect. Furthermore, the power output in this region cannot be modeled as constant as the power output usually drops in this region. Hence modeling of power curves from stall regulated wind turbines in more difficult and a larger prediction error is expected.

One explanation for the poorer performance of the models set up by [7] and [10] can be the fact that they used smaller data sets for their experiments. In both studies, data sets containing 3460 observations are used from which 2589 observations were used for training and 871 for validation (compare values in Table III). Furthermore, the settings for the NN structure in [7] and [10] appear to be not ideally set, since the performance of the k-NN model is found to be better, which is not consistent with the results obtained in this paper.

2) Performance Based on Abnormal Power Output Identification: Fig. 10(a) shows the prediction error of the ANFIS model over time and Fig. 10(b) the corresponding prediction error probability distribution for the validation data set of WEC A.

In [15], it is shown that prediction error averaging can drastically decrease the variance of the prediction and thus increase the fault visibility by allowing tighter alarm limits. This is because of the prediction errors of successfully trained models usually being normally distributed [compare Fig. 10(b)]. Depending on the application, evaluating the turbine performance

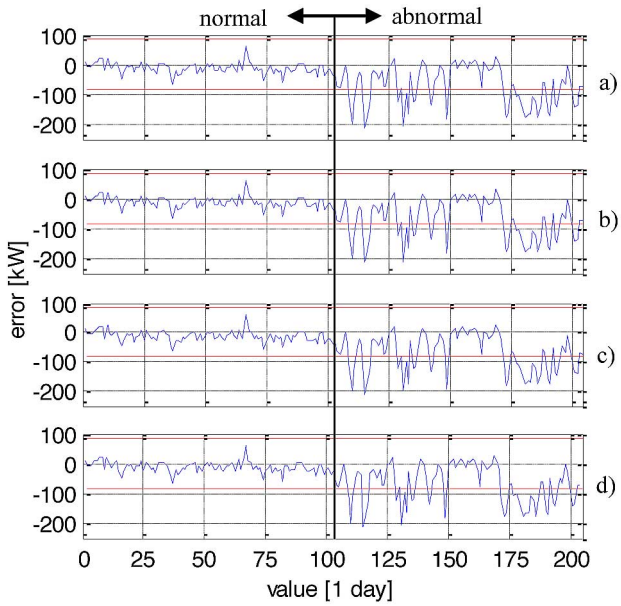

Fig. 11. Averaged prediction error of the different models in case of normal and abnormal turbine power performance (WEC A). (a) CCFL; (b) NN; (c) k-NN; (d) ANFIS.

TABLE IV
TYPE 2 MODEL PERFORMANCE COMPARISON

	Un-scaled	Scaled to 0…100kW			Normalized
	MAE [kW]	MAE [kW]	RMS [kW]	SD [kW]	MAPE [%]
CCFL	33.53	1.68	2.42	2.41	8.91
NN	32.49	1.62	2.34	2.34	8.38
k-NN	85.35	4.27	6.02	5.96	25.48
ANFIS	32.01	1.60	2.30	2.30	8.25

once a day can be sufficient. For those applications, averaging is recommended.

Fig. 11 shows the performance of the different models both for normal turbine behavior and a period where the turbine power output is lower than expected due to a wrong controller setting after a software update. Red lines mark the normal operational range being three times the standard deviation of the prediction error in the normal period.

All models show similar performances and are capable in detecting the decrease of the turbine performance. For all models, the prediction error leaves the normal performance range, on day 109 of the considered time series.

B. Type 2 Model Performance Comparison

1) Performance Based on MAE, RMS, MAPE, and SD: The performances of the different models are summarized in Table IV. Again, the given values are averages of the three turbines considered.

The ANFIS models and the NN models perform best. Both model types are known to be well suited for nonlinear modeling due to the large flexibility when given enough neurons/MFs. The k-NN approach can also be used to model nonlinear relations as it was shown earlier. However, the existing data is organized in a d-dimensional space, where d is the dimension of the input space, i.e., the number of variables. By adding two more input variables, the number of considered neighbor's k must decrease (compare Fig. 7). As a consequence, the predictions also become more sensitive to outliers in the reference data set due to the reduced number of considered and existing neighbors. In fact, the distances between the neighbors may increase since the data is spread among a larger space. Hence the number of observations should be larger in order to achieve better model performance. The CCFL algorithm also organizes the data in a d dimensional space. Depending on the number of cluster centers and radii set, the cluster centers are identified for each dimension separately. For each cluster center, an MF is initialized and linked via fuzzy logic rules. The MF shapes are Gaussian normal distributions. The performance of the CCFL is somewhat lower than the performance of the NN and ANFIS, which is potentially caused by the MF shapes.

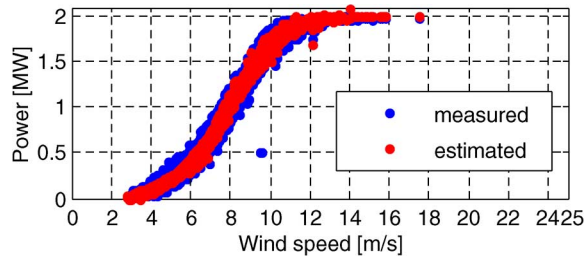

Fig. 12. Power curve of WEC A for the validation period measured and estimated by the ANFIS model (type 2).

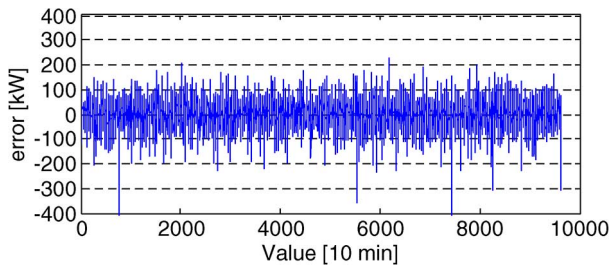

Fig. 13. Prediction error of the ANFIS model for the validation data set (WEC A; type 2).

The measured and estimated power curve of the validation data set from WEC A is visualized in Fig. 12 using the ANFIS model.

In comparison to Fig. 9, the modeled power curve has a broader scatter and some of the variance in the measurements is modeled, and incorporated in the model.

2) Performance Based on Abnormal Power Output Identification: Fig. 13 shows the prediction error of the ANFIS model for the validation data set of WEC A. The corresponding prediction error probability distribution is compared to the distribution of type one in Fig. 14.

It is important to highlight here, that when moving from model type one to model type two the prediction error magnitudes decreased.

The probability density distribution becomes slender and the probability of obtaining a low prediction error increases. The capability of detecting abnormal turbine performance of the different models is visualized in Fig. 15, using the same definition

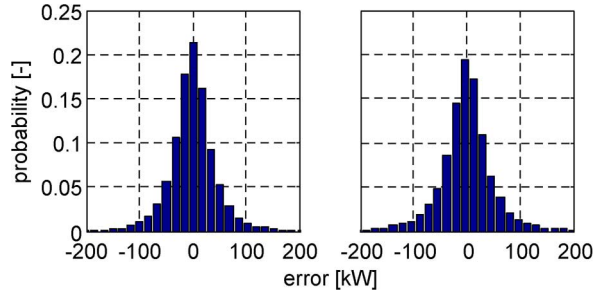

Fig. 14. Probability distribution of the prediction error (WEC A). Left: type 2; right: type 1.

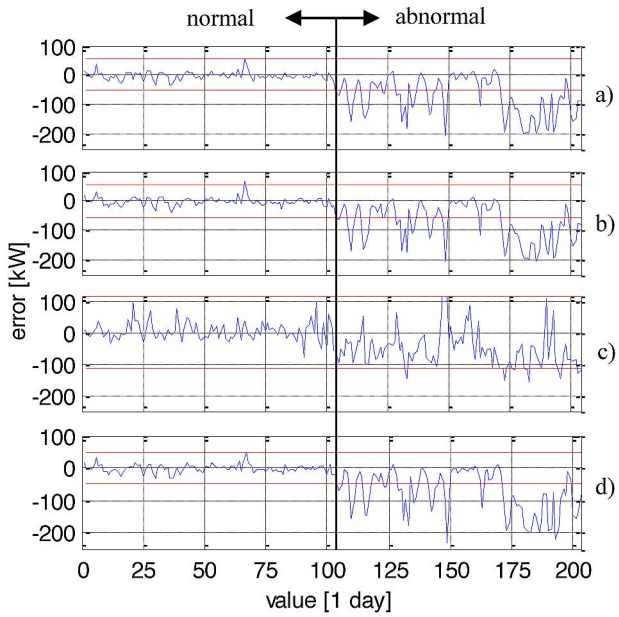

Fig. 15. Averaged prediction error of the different models in case of normal and abnormal turbine power performance (WEC A). (a) CCFL; (b) NN; (c) k-NN; (d) ANFIS.

TABLE V
ANOMALY DETECTION CAPABILITIES OF THE DIFFERENT MODELS

	Day of NR violation	Model deviation [kW]	NR [kW]
CCFL	104	-54.98	±53.60
NN	105	-64.19	±54.99
k-NN	132	-124.50	±115.10
ANFIS	104	-54.57	±50.17

of the normal operational range as in the previous section and the same time series.

Table V gives an overview about the determined normal ranges (NRs) for the different models and states the day the NR is violated, together with the model deviation at this day.

The performance of the k-NN model is poor. Not only is the SD of the prediction error large, but also the anomaly is detected late in comparison to the other models. Based on the day of NR violation, the ANFIS and the CCFL model perform equally well and highlight the anomaly one day earlier than the NN model. Considering also the difference between the NR and the model deviation at the point of NR violation as an indicator for the fault visibility, the ANFIS model performs better with a difference of -4.40 kW in comparison to -1.38 kW for the CCFL model.

In comparison to model type one, the consideration of the ambient temperature and the wind direction in model type two lead to more accurate models, i.e., lower variance in the prediction error and enabled anomaly detection five days earlier.

VI. CONCLUSION

Next to the wind speed, the ambient temperature and the wind direction were proven to be important parameters when setting up data-mining models for wind turbine power curve monitoring. The results show that when only the wind speed is used as input, the performance of the tested models is of comparable magnitude. Using data from pitch regulated modern wind turbines, all models set up perform better than the ones described in the literature. No major differences exist in the ability to detect abnormal turbine performance.

The results further show that when enhancing the CCFL, NN, and the ANFIS models by using the ambient temperature and the wind direction as additional inputs, earlier detection of abnormal turbine performance is possible due to reduced variance in the prediction errors. For the available data set, the anomaly is detected up to five days earlier than with the models using only the wind speed.

Using the enhanced models, the differences between ANFIS, CCFL, and the NN in RMS, MAE, MAPE, and SD are rather small. The ANFIS model not only showed the best performance in terms of this metric, but also in terms of abnormal power output detection, whereas the k-NN model performed worst. One explanation for the poor performance of the k-NN model can be the higher dimensionality of the space where the data are distributed when adding more input variables. By adding more input variables, the number of considered neighbors must decrease. As a consequence, the predictions become more sensible to outliers in the reference data set.

Depending on the wind resources at the turbine site, it can take a longer time before data up to the cut-out speed are collected. Hence the available data sets for model training may not include all wind speeds. For pitch regulated wind turbines, this does not endanger the success of the proposed procedure. In case no high wind speed measurements are available, the power curve characteristics allow modeling the power output in this region as constant rated power. Thus the power curve can be split into two parts: the first is modeled by the set up model and the second is constant rated power for high wind speeds up to the cut-out speed.

Even with the enhanced models, it still remains a challenge to identify small performance deviations in the 10-min average period. Future research should, therefore, focus on how to further reduce the variance in the prediction errors. This could for instance be achieved by considering the air pressure or other parameters influencing the turbine power output.

In order to avoid financial loss, turbine operators are strongly recommended to monitor the performance of their turbines with either of the methods proposed. Ideally this would be done using the enhanced ANFIS models. Those models not only showed

best performance but also their results can be, if necessary, backtracked and both the rules and the MF can be added or further modified by an expert to fine-tune them.

In case no advanced modeling tools are available which would allow ANFIS models to be set up, easy to program methods such as the k-NN can be used for power curve monitoring when considering only the wind speed. The drawback is, however, a larger prediction error and delayed anomaly detection.

Once performance decrease is detected, operators can take action to improve the performance, dependent on the root cause identified (here the problem was fixed by a further software update).

REFERENCES

[1] M. Wächter, P. Milan, T. Mücke, and J. Peinke, "Power performance of wind energy converters characterized as stochastic process: Applications of the Langevin power curve," *Wind Energy*, vol. 14, Wind Energy, pp. 711–717, 2011.

[2] S. Li, E. O'Hair, and M. G. Giesselmann, "Using neural networks to predict wind power generation," in *Proc. Int. Solar Energy Conf.*, Washington, DC, 1997, pp. 415–420.

[3] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998, ISBN: 0132733501.

[4] S. Li, D. C. Wunsch, E. O'Hair, and M. G. Giesselmann, "Comparative analysis of regression and artificial neural network models for wind turbine power curve estimation," *J. Solar Energy Eng.*, vol. 123, pp. 327–331, 2001.

[5] T. Üstüntaş and A. Duran Şahin, "Wind turbine power curve estimation based on cluster center fuzzy logic modeling," *J. Wind Eng. Ind. Aerodyn.*, vol. 96, pp. 611–620, 2008.

[6] B. Stephen, S. J. Galloway, D. McMillan, D. C. Hill, and D. G. Infield, "A Copula model of wind turbine performance," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 965–966, May 2011.

[7] A. Kusiak, H. Zheng, and Z. Song, "Models for monitoring wind farm power," *Renew. Energy*, vol. 34, pp. 583–590, 2009.

[8] E. Frank, Y. Wang, S. Inglis, G. Holmes, and I. H. Witten, "Using model trees for classification," *Mach. Learning*, vol. 32, no. 1, pp. 63–76, 1998.

[9] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learing Tools and Techniques*. San Mateo, CA, USA: Morgan Kaufmann, 1999, ISBN: 978-1-558-60552-7.

[10] A. Kusiak, H. Zheng, and Z. Song, "On-line monitoring of power curves," *Renew. Energy*, vol. 34, pp. 1487–1493, 2009.

[11] J.-S. R. Jang, "ANFIS: Adaptive-Newtork-Based Fuzzy Interference System," *Trans. Syst. Man Cybern.*, vol. 23, no. 3, pp. 665–685, 1993.

[12] A. Kusiak and X. Wei, "Prediction of methane production in wastewater treatment facility: A data-mining approach," *Annals Operat. Res.*, vol. 2011, pp. 1–11, 2011.

[13] M. O. L. Hansen, *Aerodynamics of Wind Turbines*. London, U.K.: Earthscan, 2008, 978-1-84407-438-9.

[14] *Wind Turbines—Part 12-1: Power Perfromance Measurements of Electricity Producing Wind Turbines*, International Standard, IEC 61400-12-1, International Electrotechnical Commision, 2005.

[15] M. Schlechtingen and I. F. Santos, "Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection," *Mech. Syst. Signal Process.*, vol. 25, no. 5, pp. 1849–1875, 2011.

[16] S. L. Chiu, "Fuzzy model identification based on cluster estimation," *J. Intelligent Fuzzy Syst.*, vol. 2, pp. 267–278, 1994.

[17] T. J. Ross, *Fuzzy Logic With Engineering Applications*, 3rd ed. Hoboken, NJ, USA: Wiley, 2010.

[18] L. Tarassenko, *Guide to Neural Computing Applications*. Amsterdam, The Netherlands: Elsevier, 1998, ISBN: 0340705892.

[19] M. Y. Rafiq, G. Bugmann, and D. J. Easterbrook, "Neural network design for engineering applications," *Comput. Structures*, vol. 79, pp. 1541–1552, 2001.

[20] K. Swingler, *Applying Neural Networks—A Practical Guide*. Amsterdam, The Netherlands: Elsevier, 1996, ISBN: 0340705892.

[21] L. Tarassenko, *Guide to Neural Computing Applications*. Amsterdam, The Netherlands: Elsevier, 1998, ISBN: 0340705892.

[22] C. M. Bishop, *Pattern Recognition and Machine Learing*. New York, NY, USA: Springer, 2009, ISBN: 978-0-387-31073.

[23] E. H. Mamdani, "Application of fuzzy logic to approximate reasoning using linguistic synthesis," *Trans. Comput.*, vol. 23, no. 12, pp. 1182–1191, 1977.

[24] M. Sugeno, *Industrial Applications of Fuzzy Control*. Amsterdam, The Netherlands: Elsevier Science, 1985.

[25] M. Schlechtingen, I. F. Santos, and S. Achiche, "Wind turbine condition monitoring based on SCADA data using normal behavior models. Part 1: System description," *J. Appl. Soft Comput.*, vol. 13, pp. 259–270, 2013.

Meik Schlechtingen received a diploma from the University of Applied Science Gießen-Friedberg, Germany, in 2008, and the M.Sc. degree in engineering from the Technical University of Denmark, Denmark, in 2010. Since 2010, he has been working toward the Ph.D. degree at the Technical University of Denmark.

He has been employed by the EnBW Erneuerbare Energien GmbH since 2010. His research focuses on wind turbine condition monitoring using SCADA data and vibration measurements. Prior to his Ph.D. study, he worked at a wind turbine manufacturer in the technical support department.

Ilmar Ferreira Santos received the Dr.-Ing. degree from the Technical University of Munich, in Germany, the dr. techn. degree from Technical University of Denmark, and the livre-docente degree from State University of Campinas, Brazil.

He is full professor at Technical University of Denmark, Department of Mechanical Engineering. He works in the field of multiphysics design, optimization, monitoring, and control of electro-mechanical machine elements. His research interests focus upon machinery dynamics, tribology, control, and mechatronics.

Sofiane Achiche received the M.Sc.A. and Ph.D. degrees from École Polytechnique de Montréal, Canada.

He is a professor with École Polytechnique de Montréal, Mechanical Engineering Department, Design of Machinery Section. His research interests focus upon evolutionary computational intelligence applied to engineering problems such as condition monitoring. Furthermore, he works in the field of mechatronics design as well as understanding and modeling activities of new product development processes for decision support purposes.