IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 7, 2021, accepted June 14, 2021, date of publication June 18, 2021, date of current version June 30, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3090434

A Hybrid Intelligent Model for the Condition Monitoring and Diagnostics of Wind Turbines Gearbox

AZIM HEYDARI[1], (Student Member, IEEE), DAVIDE ASTIASO GARCIA[2], AFEF FEKIH[3], (Senior Member, IEEE), FARSHID KEYNIA[4], LINA BERTLING TJERNBERG[5], (Senior Member, IEEE), AND LIVIO DE SANTOLI[1]

[1]Department of Astronautics, Electrical and Energetic Engineering (DIAEE), Sapienza University of Rome, 00184 Rome, Italy
[2]Department of Planning, Design, and Technology of Architecture, Sapienza University of Rome, 00185 Roma, Italy
[3]Department of Electrical and Computer Engineering, University of Louisiana at Lafayette, Lafayette, LA 70504, USA
[4]Department of Energy Management and Optimization, Institute of Science and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman 76176195892, Iran
[5]School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology Stockholm, 100 44 Stockholm, Sweden

Corresponding author: Afef Fekih (afef.fekih@louisiana.edu)

ABSTRACT Wind turbines (WTs) are often operated in harsh and remote environments, thus making them more prone to faults and costly repairs. Additionally, the recent surge in wind farm installations have resulted in a dramatic increase in wind turbine data. Intelligent condition monitoring and fault warning systems are crucial to improving the efficiency and operation of wind farms and reducing maintenance costs. Gearbox is the major component that leads to turbine downtime. Its failures are mainly caused by the gearbox bearings. Devising condition monitoring approaches for the gearbox bearings is an effective predictive maintenance measure that can reduce downtime and cut maintenance cost. In this paper, we propose a hybrid intelligent condition monitoring and fault warning system for wind turbine's gearbox. The proposed framework encompasses the following: a) clustering filter- (based on power, rotor speed, blade pitch angle, and wind speed signals)-using the automatic clustering model and ant bee colony optimization algorithm (ABC), b) prediction of gearbox bearing temperature and lubrication oil temperature signals- using variational mode decomposition (VMD), group method of data handling (GMDH) network, and multi-verse optimization (MVO) algorithm, and c) anomaly detection based on the Mahalanobis distances and wavelet transform denoising approach. The proposed condition monitoring system was evaluated using 10 min average SCADA datasets of two 2 MW on-shore wind turbines located in the south of Sweden. The results showed that this strategy can diagnose potential anomalies prior to failure and inhibit reporting alarms in healthy operations.

INDEX TERMS Automatic clustering, condition monitoring, forecasting, GMDH neural network, multi-verse optimization, wind turbine assessment.

I. INTRODUCTION

Wind energy is currently widely used in several countries as a clean, cost-effective and sustainable source of renewable energy [1]. Wind turbines' operation in harsh environment and in the presence of highly variant stochastic loads, however, makes them prone to sensor, actuator and component faults, thereby requiring increased frequency of planned maintenance scheduling [2], [3]. This latter, however, leads to higher maintenance costs and increased downtime and subsequently reduced power production. To lower the cost of maintenance, decrease downtime and improve wind turbine's reliability, in the presence of faults, various condition monitoring techniques based on data obtained by the wind turbine's Supervisory Control And Data Acquisition (SCADA) system have been proposed in the literature [4]–[7].

Data-driven methods were recently shown to be quite effective in condition monitoring [8], [9]. In [6], Bangalore and Bertling Tjernberg, introduced an artificial neural network (ANN) condition monitoring model according

The associate editor coordinating the review of this manuscript and approving it for publication was Dazhong Ma.

to data obtained from the SCADA system. The proposed model was used for gearbox bearings with actual data from onshore wind turbines [10]. In [11], Cheng *et al.* developed a novel method with a doubly fed induction generator (DFIG) stator current signal to diagnose faults in the wind turbine drivetrain gearbox under nonstationary conditions. In [12], Dao, *et al.*, proposed a condition monitoring and failure detection model based on the co-integration analysis of SCADA data. This approach was able to appropriately analyze nonlinear data trends, constantly monitor the wind turbine, and reliably diagnose abnormal conditions. In [13], Sun *et al.*, proposed a model for wind turbine anomaly detection based on various wind turbine condition parameter prediction approaches and a fuzzy theory model. In [14], Sanchez *et al.* developed a technique for fault detection for WT considering model parametric uncertainties and noise according to interval observers and analytical redundancy relations. Zhang *et al.*, presented a model to identify WT state parameters anomalies valid for condition parameters ranges fluctuating within the SCADA alarm threshold [15]. Qu *et al.* proposed a WT fault diagnosis technique with SCADA data according to the expanded linguistic terms and rules through non-singleton fuzzy logic [16]. In [17], deep neural network (DNN)-based framework was considered to detect WT gearbox faults. [18] developed a fault diagnosis system based on adaptive neuro-fuzzy inference system and hybrid models. A multi-scale convolutional neural networks-based fault diagnosis method was introduced in [19] for gearbox health monitoring.

Artificial intelligent approaches and deep learning techniques were recently introduced to automatically make timely decisions on the running health of wind turbines based on massive data sets. Intelligent fault diagnosis typically includes the following three steps: signal acquisition, feature extraction, and fault recognition based on techniques such as statistical learning theories, intelligent signal processing and artificial intelligence techniques [17], [20]–[22]. Ongoing research studies have shown that deep learning approaches yield better efficiency and accuracy in monitoring the operating conditions of the turbine. Ref [23] utilized a linear support vector machine to detect wind turbines' faults. A hierarchical event detection method based on spectral theory of multidimensional matrix was proposed in [24] for the fault detection of a power systems using massive data. A gearbox fault diagnosis approach based on a novel hybrid feature reduction approach was proposed in [25]. This approach mixed the optimization objectives of the principal component analysis (PCA) and locally linear embedding (LLE) to identify a mapping that simultaneously responds to the optimization objectives of PCA and LLE. A tacho-less order tracking technique was proposed in [26] to identifying WT gearbox faults in non-stationary conditions, without the need for conventional instantaneous angular speed (IAS) calculations. Risk management techniques were considered in [27] to evaluate the effects of some risk factors that affect the energy production of a wind farm. A data-driven model was proposed in [28] to assess the function of wind turbines at past and future time intervals. Parametric Copula models were considered in data [29] to accurately assess the performance of wind turbines based on real data sets. An intelligent SCADA data-driven, nonparametric approach was proposed in [30] for wind turbine condition monitoring. The approach applied the Gaussian process and regression tree techniques to calculate the power curve of a wind turbine and subsequently determine functional anomalies based on a comparative analysis. GP and regression models were developed using evolutionary strategy algorithms [30]. To analyze the influence of wind turbine operational variables on the precision of the model and its uncertainty, a Gaussian Process (GP) was provided in [31]. The findings indicate that considering functional parameters can enhance the performance of the GP model precision and eliminate the uncertainty in forecasting the power curve.

Although intelligent methods have recently introduced to WT condition monitoring, to the best of our knowledge, none have considered a hybrid intelligent approach nor a prediction model for parameter forecasting. These latter have the ability to fully handle the uncertainties present in large data sets, thus yielding better performance in terms of fault detection, feature selection, signal decomposition, clustering, and forecasting.

In this paper, we propose a hybrid intelligent condition monitoring system for wind turbine's gearbox. Its main contributions are as follows:

- An ant bee colony algorithm (ABC)-based hybrid automatic clustering filter model to cluster the signals (i.e. wind speed, power production, rotor speed, and pitch blade angle) affecting the gearbox performance.
- A deep learning prediction model based on GMDH neural network and multi-verse optimization algorithm for bearing temperature and lubrication oil temperature forecasting.
- An anomaly detection strategy based on the Mahalanobis distance calculation and wavelet transform de-noising method to detect possible anomalies and prevent failure occurrence.
- Practical implementation of the proposed combined deep learning model to the real SCADA data of two on-shore wind turbines located in the south of Sweden.

The remainder of this paper is organized as follows. Some preliminaries are provided in section II. The proposed anomaly detection model is detailed in section III. The experimented results are given in section IV. Finally, some concluding remarks are provided in section V.

II. Preliminaries

A. WAVELET TRANSFORM

There are two groups of wavelet transforms: continuous wavelet transform (*CWT*) and discrete wavelet transform (*DWT*). The CWT $W_{(a,b)}$ of signal $f_{(x)}$ considering a wavelet

$\varphi_{(x)}$ was proposed in [32], [33]:

$$W_{(f,a,b)} = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f_{(x)} \phi_{\left(\frac{x-b}{a}\right)} dx \tag{1}$$

In which a determines the wavelet spread and b characterizes its central location. $\varphi_{(x)}$ represents the mother wavelet. A $W_{(a,b)}$ coefficient indicates the extent to which the scaled/translated mother wavelet and the main signal f(x) are matched. Therefore, $W_{(a,b)}$ as the wavelet coefficients, which are related to a specific signal, are the signal wavelet indicator for the main wavelet. Because *CWT* can be obtained when the mother wavelet is continuously scaled and translated, substantial redundant data is produced. Thus, the scaling and translating of the mother wavelet may also be done by specific scales and positions commonly on the basis of powers of two or DWT [34]. Such technique was shown to be more effective than the *CWT* [32]. Thee *DWT* of a signal $f_{(t)}$ is defined by:

$$W_{(m,n)} = 2^{-(m/2)} \sum_{t=0}^{T-1} f_{(t)} \varphi_{\left(\frac{t-n.2^m}{2^m}\right)} \tag{2}$$

where T represents the length of signal $f(x)$. The scaling and translation factors are functions of m and n as integer variables ($a = 2m$, and $b = n.2m$); t indicates the discrete time index.

B. VARIATIONAL MODE DECOMPOSITION

Dragomiretskiy and Zosso developed a Variational mode decomposition (VMD) model as a novel signal decomposition technique [35]. VMD provides better performance in terms of sampling and noise than available methods, such as EMD [35]. It is a complete non-recursive VMD model, in which the modes are extracted at the same time. It follows a group of modes along with their center frequencies in a respective order. Hence, the modes collectively reproduce the input signal, and each mode is smoothly following the baseband demodulation [35].

$$\min_{u_k, w_k} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-jw_k t} \right\|_2^2 \right\}$$
$$s.t. \sum_k u_k = f(t) \tag{3}$$

The mode u_k should be mainly compact near the center pulsation w_k [35]. The VMD enables estimating the bandwidth through the H^1 Gaussian smoothness of the demodulated signal, i.e. the squared L2-norm of the gradient [35]. The set of K modes as well as their center frequencies are respectively presented with u_k and w_k. Moreover, the Dirac distribution is shown with δ, the time script $j^2 = -1$ is presented with t, and the convolution operator is represented by $*$. The quadratic penalty along with λ Lagrangian multipliers was expressed to create an unconditional problem. The following completed Lagrangian can be provided [35]:

$$L(u_k.w_k, \lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times u_k(t) \right] \times e^{-jw_k t} \right\|_2^2 + \left\| f - \sum_k u_k \right\|_2^2 + \left\langle \lambda, f - \sum_k u_k \right\rangle, \tag{4}$$

where the balancing factor related to the data-fidelity constraint is indicated with α. The alternate direction method of multipliers can solve Eq. (4). The mode $u_k(w)$ in the frequency domain is indicated using Eq. (5), the center frequencies w_k are demonstrated using Eq. (6), and Eq. (7) is used to update λ. The mode $u_k(t)$ in the time domain can be achieved as the real part in the inverse Fourier transform of $u_k(w)$ using Eq. (5) [35]:

$$\hat{u}_k^{n+1} = \frac{\hat{f}(w) - \sum_{i \neq k} \hat{u}_i(w) + \frac{\hat{\lambda}(w)}{2}}{1 + 2\alpha (w - w_k)^2} \tag{5}$$

$$w_k^{n+1} = \frac{\int_0^\infty w \left| \hat{u}_k(w) \right|^2 dw}{\int_0^\infty \left| \hat{u}_k(w) \right|^2 dw} \tag{6}$$

$$\hat{\lambda}^{n+1}(w) = \hat{\lambda}^n(w) + \tau \left(\hat{f}(w) - \sum_k \left| \hat{u}_i(w) \right| \right) \tag{7}$$

C. INTELLIGENT AUTOMATIC CLUSTERING

Cluster validity is associated with the statistical–mathematical functions applied to quantitatively evaluate the clustering algorithm findings. In general, the cluster validity functions have two goals. 1) determine the clusters' numbers, 2) indicate the related best partition. The best measure for validity is obtained by repeatedly running the algorithm using varied classes as input followed by selecting the data partitioning [36]. A validity function is often needed to consider the two partitioning features: a) Cohesion: patterns in a cluster need to highly resemble one another. Patterns' fitness variance in a cluster can represent cohesion and compactness in the cluster, b) Separation: separation of clusters is crucial. The space between the cluster centers (Euclidean distance) indicates cluster separation [37].

In terms of crisp clustering, Dunn's index (DI) [38], Calinski–Harabasz criterion [39], DB index [40], Pakhira Bandyopadhyay Maulik (PBM) index [41], and the CS measure [42] are the most popular indexes. They are optimizing in nature and accurately yield the proper partitions. Due to their optimizing feature, cluster validity indices were shown to be the best tools along with the optimization algorithms, such as PSO, GA, TS, etc. In this paper, crisp clustering is performed using the Davies-Bouldin (DB) function.

DB Function: It is regarded as a function of the sum of the ratio of within-cluster scatters and between-cluster separation, which employs clusters and their sample mean. The within i^{th} cluster scatter is initially defined followed by the i^{th}

and j^{th} clusters,

$$S_{i,q} = \left[\frac{1}{N_i} \sum_{\vec{X} \in C_i} \left\| \vec{X} - \vec{m}_i \right\|_2^q \right]^{\frac{1}{q}} \tag{8}$$

$$d_{i,j,t} = \left\{ \sum_{p=1}^{d} \left| m_{i,p} - m_{j,p} \right|^t \right\}^{\frac{1}{t}} = \left\| \vec{m}_i - \vec{m}_j \right\|_t \tag{9}$$

where $\vec{m}_i$ represents the i^{th} cluster center, q indicates an integer, q and t are separately selectable, Ni shows the elements numbers within the i^{th} cluster C_i.

$R_{i,qt}$ is then expressed as:

$$R_{i,qt} = \max_{j \in K, i \neq i} \left\{ \frac{S_{i,q} + S_{j.q}}{d_{i,j,t}} \right\} \tag{10}$$

Ultimately, the DB measure is determined as follows:

$$DB\ (K) = \frac{1}{K} \sum_{i=1}^{K} R_{i,qt}. \tag{11}$$

The smallest $DB(K)$ index shows a valid optimal partition.

D. ARTIFICIAL BEE COLONY ALGORITHM (ABC)

Karaboga designed the ABC algorithm in 2005 [43] to improve numerical function optimization with respect to cooperative foraging and waggle-dancing of honey bees. Detection of an optimal ABC solution is similar to the foraging process in bees. The location of the source is regarded as a possible solution, and the nectar quantity in each source is indicative of its fitness. Artificial bees can be classified into employed and unemployed bees (i.e., onlooker and scout). The scout and onlooker bees each occupy half of the colony. As every food source is linked to one employed bee, the number of employed bees represents the number of source positions (solutions). There are four different phases in the ABC algorithm: initialization, employed bee, onlooker bee, and scout bee.

E. GMDH NEURAL NETWORK

A.G. Ivakhnenko developed GMDH as a heuristic self-organizing approach employed by complex nonlinear systems [44]. It is an algorithm according to the self-organizing data mining of the external criterion, in which the Volterra–Kolmogorov–Gabor (VKG) polynomial indicates the association of input and output factors in a network characterized by different inputs and single output [44]:

$$\hat{y} = a_0 + \sum_{i=1}^{m} a_i x_i + \sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij} x_i x_j + \sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{k=1}^{m} a_{ijk} x_i x_j x_k + \ldots \tag{12}$$

where the input variables and the unclear coefficients are respectively presented with $x_1, x_2, \ldots, x_m$ and $a_0, a_1, \ldots, a_{ijk}$, and the number of input variables is shown with m. For the majority of cases, the VKG series general equation can be simplified as a polynomial with second-order relationship and two variables [45]:

$$\hat{y} = \hat{f}\ (x_1, x_2) = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1 x_2 + a_4 x_1^2 + a_5 x_2^2 \tag{13}$$

An iterative method including training/testing phases can construct using the Group Method of Data Handling (GMDH) structure. Through training of a network, unclear quadratic polynomial parameters can be determined by reducing the number of errors between the model estimated data and the experimental values:

$$Min \left(\sum_{i=1}^{N_t} \left[\hat{y}_i - y_i \right]^2 = \sum_{i=1}^{N_t} \left[\hat{f}\left(x_{ip}, x_{iq} \right) - y_i \right]^2 \right), \tag{14}$$

where N_t represents the training value number. In the testing phase, the best combination of variables is chosen through testing values [46]–[48]. The architecture of the GMDH deep learning is depicted in Figure 1.

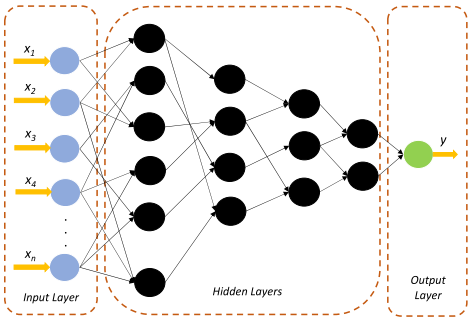

FIGURE 1. The structure of GMDH deep learning NN.

F. MULTI-VERSE OPTIMIZATION ALGORITHM

Seyedali Mirjalili *et al.* proposed the Multi-Verse Optimization (MVO), as a nature-inspired heuristic optimization algorithm in 2016 [49] inspired by the multiverse theory in astrophysics. It uses 3 concepts in astrophysics, such as white, black, and wormholes studying the universe evolution. While presenting the MVO algorithm the terms are used as follows: a solution is provided by the universe, an object is related to a solution, generation/iteration can be indicated with time, and a universe objective is shown with the inflation rate.

Regarding MVO, every solution equals a universe with potential white, black, and worm holes. For improving each solution quality, it is more probable for matter emitters (i.e., white holes) to be indicated in a solution characterized by more proper objective value. Conversely, matter attractors (black holes) are seen in a solution characterized by the worse objective value. Therefore, values related to good solution variables are transmitted to poor solutions, which can improve poor solutions leading to the improvement of the mean objective value from all solutions. The principal

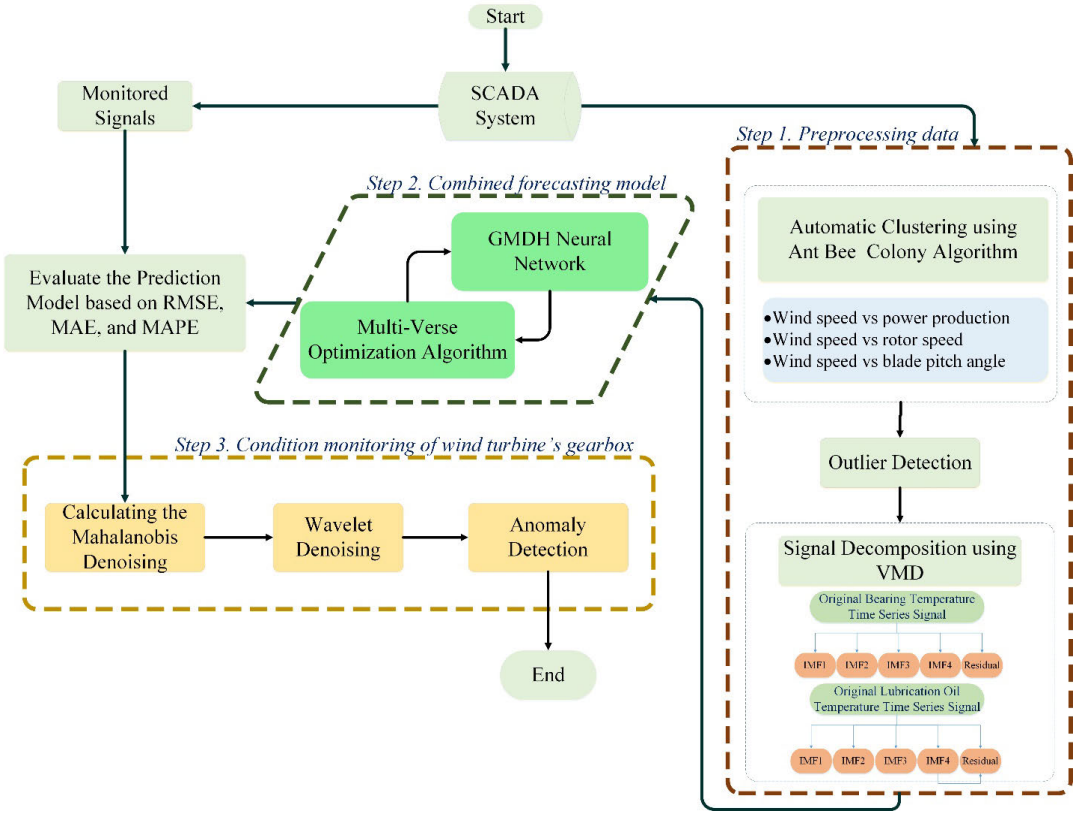

FIGURE 2. Framework of the proposed condition monitoring model.

mathematical model related to the algorithm is associated with Eqs. (15) and (16):

$$x_i^j = \begin{cases} x_k^j & r_1 < NI\,(U_i) \\ x_i^j & r_1 \geq NI\,(U_i) \end{cases} \tag{15}$$

x_i^j indicates the j^{th} object related to the i^{th} universe, r_1 represents a random number between 0 and 1, $NI(U_i)$ indicates the normal inflation rate related to the i^{th} universe, and x_k^j shows the j^{th} object related to the k^{th} universe.

$$x_i^j = \begin{cases} x_j + TDR \times ((Ub_b - Lb_b) \times r_4 \times Lb_b) \\ \qquad r_3 < 0.5 \;\; \& \;\; r_2 < WEP \\ x_j - TDR \times ((Ub_b - Lb_b) \times r_4 \times Lb_b) \\ \qquad r_3 \geq 0.5 \;\; \& \;\; r_2 < WEP \\ x_i^j \quad r_2 \geq WEP \end{cases} \tag{16}$$

where x_j represents the j^{th} centroid related to the most appropriate universe achieved, Ub_b and Lb_b present the upper and lower bounds respectively, Traveling Distance Rate (TDR) and Wormhole Existence Probability (WEP) are coefficients, r_2, r_3 and r_4 are random numbers between 0 and 1. WEP indicates the potential existence of wormhole in the universes. Through iterations, it experiences a linear increase to confirm the exploitation. TDR indicates that an object moves through wormhole all over the best universe. TDR experiences an increase during algorithm to achieve clear exploitation all over the best universe. The following equations indicate both WEP and TDR:

$$WEP = Min + Iteration \times \left(\frac{Max - Min}{L}\right) \tag{17}$$

$$TDR = 1 - \frac{Iteration^{1/p}}{L^{1/p}} \tag{18}$$

The minimum and maximum values are demonstrated with *Min* and *Max* (0.2 and 1, respectively), the current iteration is shown with Iteration, the maximum iteration number indicates L, and the exploitation accuracy is presented with p (typically, with the value of 6). The MVO algorithm forms a group of random universes including objects aiming at transferring from a high inflation rate universe to a low inflation rate universe via white and black holes. Objects are transported in random via wormhole all over the best universe, and the process is iterated to achieve a global optimal solution [50].

III. PROPOSED ANOMALY DETECTION MODEL

The schematic of the proposed combined intelligent model is proposed for bearing fault detection of wind turbine's gearbox based on temperature signals is illustrated in Figure 2.

The proposed model consists of the following five phases:

Phase 1 (Clustering Process): In this phase, three steps are considered to process the SCADA data of wind turbines and eliminate the outliers.

a) The proposed clustering method based on automatic clustering and artificial bee colony algorithm (DB-ABC) is applied to classify the data into different clusters which indicate the various states under normal operations.
b) The input signals have been decomposed using the VMD decomposition model.
c) GMDH network has been applied as a main forecaster engine to predict the output variables.
d) In order to optimize the parameters of the GMDH network, the MVO optimization algorithm has been used.
e) If the Mahalanobis distance of a record is not in the range of 3 standard deviations from the means in each cluster, it is an outlier record, which should be deleted in the filter [51].
f) The operation data are standardized to the standard normal distribution for eliminating the scale impact.

In this paper, the data sets are classified into different clusters based the wind power production, rotor speed, blade pitch angle, and wind speed. These signals are considered since they reflect various operation states of the system components. Note that the blade pitch angle is considered as a parameter for cluster filtering because the correlation of blade pitch angel with the output variables is fairly high.

Figures 3-5 depict the power curve, rotor curve, and blade pitch curve for wind turbine *A*, respectively.

Phase 2 (Forecasting Model): In this phase, a new hybrid forecasting model based on variational mode decomposition, GMDH neural network, and Multi-Verse optimization algorithm is proposed for gearbox bearing temperature and gearbox oil lubrication temperature forecasting. First, the VMD method has been applied to decompose the two temperature signals into different frequencies (IMF1, IMF2, IMF3, and IMF4). Second, a mixed data model based on variational mode decomposition, power production, rotor speed, ambient temperature, nacelle temperature, and the original signals of gearbox bearing temperature and gearbox oil lubrication temperature (with five lagged values) has been developed and used as input parameters to increase the accuracy and stability of the forecasting model:

$$\textit{Gearbox Bearing Temperature } (GBT) \times \begin{cases} GBT_{(t-1)}, GBT_{(t-2)}, GBT_{(t-3)}, GBT_{(t-4)}, GBT_{(t-5)} \\ IMF1^{GBT}, IMF2^{GBT} IMF3^{GBT}, IMF4^{GBT} \\ \textit{Power Production} \\ \textit{Rotor Speed} \\ \textit{Ambiant Temperature} \\ \textit{Nacelle Temperature} \end{cases}$$

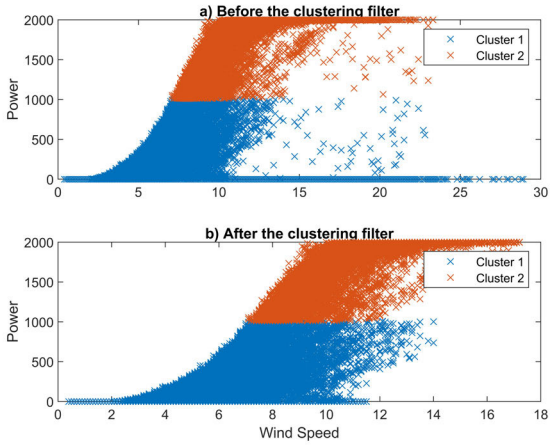

FIGURE 3. The power curve clustering for wind turbine A–a) before clustering filter and b) after clustering filter.

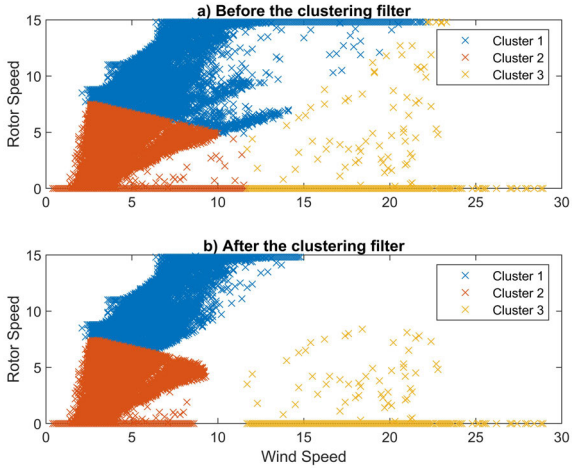

FIGURE 4. The rotor curve clustering for wind turbine A–a) before clustering filter and b) after clustering filter.

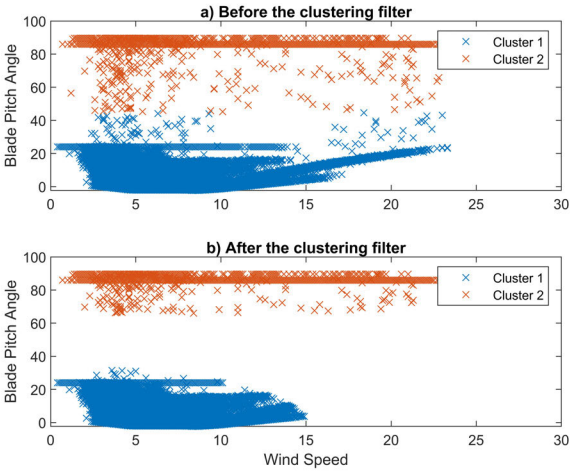

FIGURE 5. The blade pitch curve clustering for wind turbine A–a) before clustering filter and b) after clustering filter.

$$\text{Gearbox Oil Lubrication Temperature } (GOLT) \times \begin{cases} GOLT_{(t-1)}, GOLT_{(t-2)}, GOLT_{(t-3)}, GOLT_{(t-4)}, \\ \quad GOLT_{(t-5)} \\ IMF1^{GOLT}, IMF2^{GOLT} IMF3^{GBTGOLT}, IMF4^{GOLT} \\ \text{Power Production} \\ \text{Rotor Speed} \\ \text{Ambiant Temperature} \\ \text{Nacelle Temperature} \end{cases}$$

Third, the GMDH neural network applied as forecasting engine to predict the two temperature signals. Fourth, in order to obtain better forecasting results, the MVO algorithm has used to optimize the GMDH network parameters. It means, the parameters of GMDH network has been defined as an optimization problem, then the problem optimized by MVO algorithm. In addition, in evaluation part of the model, three different error criteria (RMSE, MAE, and MAPE) have been applied to assess the model performance. The mathematical equations of the error criteria are calculated as follows:

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^{m} \left(x_{true_i} - x_{predicted_i}\right)^2} \tag{19}$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left|x_{true_i} - x_{predicted_i}\right| \tag{20}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left|\frac{x_{true_i} - x_{predicted_i}}{x_{true_i}}\right| \times 100 \tag{21}$$

where x_{true} and $x_{predicted}$ are respectively the true and the predicted value. N is the number of samples. Figures 6 and 7

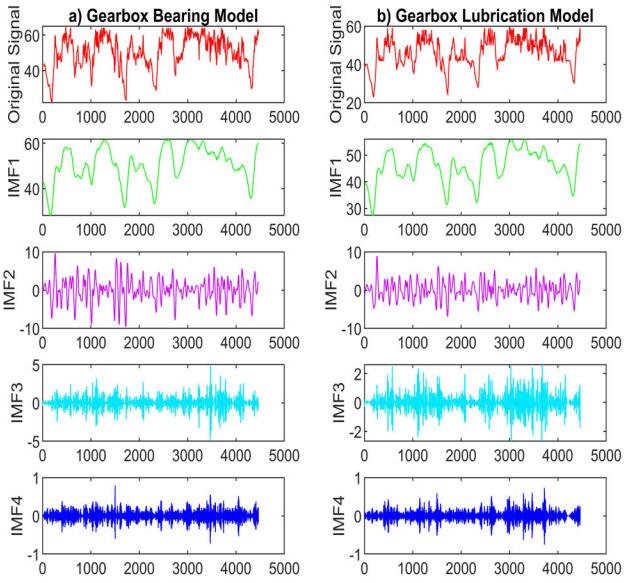

FIGURE 6. Decomposition results of the gearbox bearing temperature and the gearbox oil lubrication temperature for wind turbine A in January 2013.

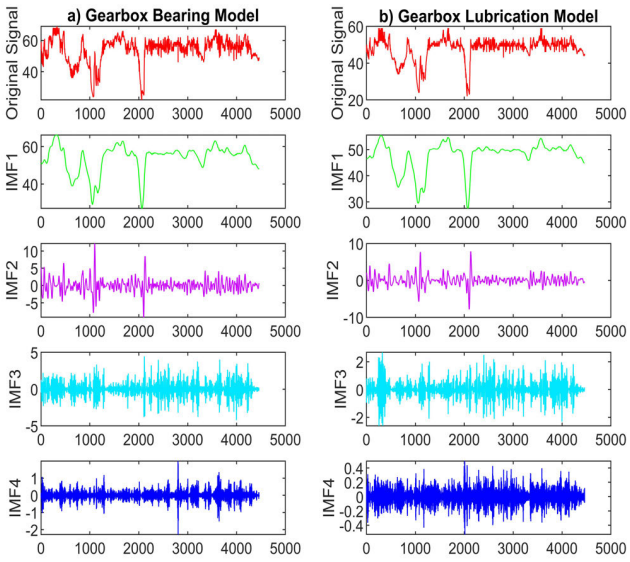

FIGURE 7. Decomposition results of the gearbox bearing temperature and the gearbox oil lubrication temperature for wind turbine B in January 2014.

presented the decomposition signals of gearbox bearing temperature and gearbox oil lubrication temperature for wind turbines *A* and *B,* respectively.

Phase 3 (Anomaly Detection Analysis): In this phase, the Mahalanobis distance is applied for assessing the deviations between the true values of the temperature signals and their forecasted values obtained from the hybrid forecasting model (VMD-GMDH-MVO). This choice is motivated by the fact that unlike the Euclidean distance, the Mahalanobis distance is scale-invariant, unit-less and considers the correlation between different variables. Consider x and y generated by the same probability distribution, their Mahalanobis distance can be determined as follows:

$$MD_i = \sqrt{(x_i - \mu)\ Cov^{-1}\ (x_i - \mu)^T} \tag{22}$$

where Cov represents the covariance matrix of x and y. While Cov represents a unit variance matrix, the Mahalanobis distance is in accordance with the Euclidean distance. In this phase, the absolute error of forecasted values and true temperature measurements are firstly calculated. Then, the Mahalanobis distances of the errors compared with others that are below the healthy states are determined as the assessed indicators for measuring the deviations between the present states and healthy operations.

In addition, the calculated Mahalanobis distances are filtered using the wavelet de-noising. They are decomposed at various levels using the wavelet base functions (including db6 at six levels denoising). The achieved detail coefficients can be determined by the threshold at each level. Next, the de-noised distances are formed according to the modified coefficients as well as the wavelet base functions. Furthermore, the upper limit of the Mahalanobis distances related to healthy states indicted by the threshold [51]. The de-noised

Mahalanobis distances in healthy operations are calculated. They are assumed to be demonstrated through a specific probability distribution. Due to the large samples, we regarded normal distributions. Then, Kolmogorov-Smirnov test was used to test the assumption and when accepted, the fitted distribution is identified and the threshold can be determined at a low probability (0.001). Otherwise, we test new probability distributions are until no rejection of the corresponding assumptions [51].

A comparison is made between the de-noised Mahalanobis distances with the obtained threshold. The WT is regarded as healthy for the distance below the threshold and the lack of anomalies in the component. If the distance value crosses the threshold it triggers a warning that shows potential operation risks; however, it is not dangerous. The alarm can be triggered in cases of continuous warnings for more than 2 h, which can warn operators about the possible beginning anomalies [51].

IV. EXPERIMENTAL RESULTS

In this paper, the SCADA data and maintenance information of two wind turbines have been used in order to evaluate the proposed model for anomaly condition monitoring of wind turbine's gearbox. The on-shore wind farm is located in the south of Sweden [52]. The results of the analysis of each turbine are separately described below.

A. FORECASTING RESULTS

Given the importance of predicting accuracy in anomaly detection analysis for condition monitoring, we consider a deep learning-based combined model to predict the following two signals: gearbox bearing temperature and oil lubrication temperature. This prediction model consists of four parts: 1) clustering, 2) signal decomposition, 3) GMDH neural network, and 4) multi-verse optimization algorithm. Table 1 indicates the prediction results of the proposed model for the gearbox bearing and lubrication temperature forecasting for two wind turbines (*Turbine-A* and *Turbine-B*).

TABLE 1. The results of the proposed prediction model for different gearbox models in two case studies.

Case study	Target Models	Models	RMSE	MAE	MAPE	R
Turbine-A	Bearing model	MLP	0.8019	0.7181	4.7822	0.9878
		VMD-GMDH-GA	0.7985	0.5871	2.2138	0.9941
		VMD-GMDH-PSO	0.7821	0.5981	1.7181	0.9923
		Proposed model	0.7611	0.5791	1.1793	0.997
	Lubrication model	MLP	0.7872	0.6716	3.6762	0.9782
		VMD-GMDH-GA	0.856	0.5342	1.7612	0.9844
		VMD-GMDH-PSO	0.8737	0.5132	1.932	0.9878
		Proposed model	0.744	0.4968	1.0694	0.9954
Turbine-B	Bearing model	MLP	0.9463	0.8721	2.659	0.9816
		VMD-GMDH-GA	0.9232	0.6812	1.4361	0.9916
		VMD-GMDH-PSO	0.9382	0.6435	1.3312	0.9935
		Proposed model	0.9157	0.6238	1.28	0.9982
	Lubrication model	MLP	0.9832	0.8712	3.0198	0.9855
		VMD-GMDH-GA	0.8738	0.5393	1.6481	0.9917
		VMD-GMDH-PSO	0.9238	0.5642	1.6512	0.9892
		Proposed model	0.9556	0.5136	1.2169	0.9963

TABLE 2. Obtained MAPE values of the proposed model for the structural analysis.

Case study	Target Models	Without VMD model	Without MVO Optimizer
Turbine-A	Bearing model	1.6618	2.0138
	Lubrication model	1.981	2.4519
Turbine-B	Bearing model	2.0163	2.201
	Lubrication model	2.1142	2.7681

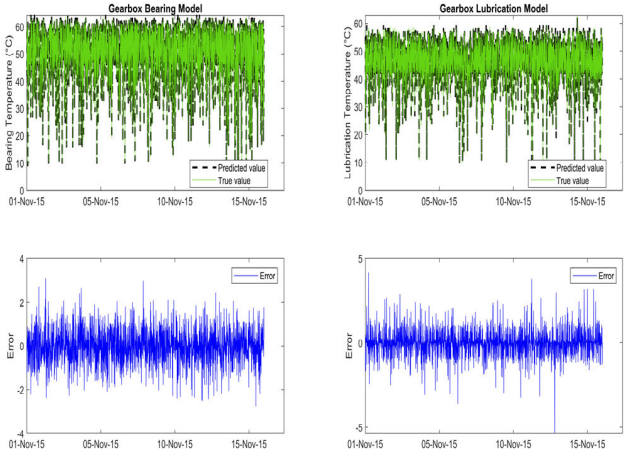

FIGURE 8. The predicted and true values of the proposed prediction model in Turbine-A: a) the gearbox bearing model, b) the gearbox lubrication model.

Figures 8 and 9 show the predicted values and true values and the error of the proposed model for the gearbox bearing temperature and lubrication oil temperature forecasting in turbines *A* and *B*, respectively.

The correlation of the predicted and true values and the error distribution histogram of the proposed prediction model in Turbines A and B are illustrated in Figures 10 and 11, respectively.

Based on these Figures (10 and 11), the correlation plots indicate the correlation between the true value and predicted value. Note that the higher this correlation, the better the

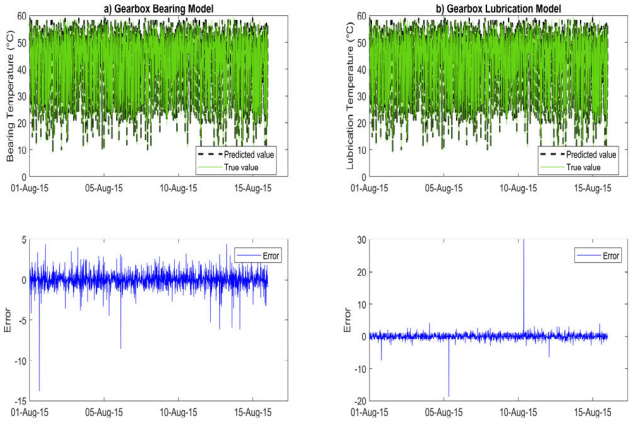

FIGURE 9. The predicted and true values of the proposed prediction model in Turbine-B: a) the gearbox bearing model, b) the gearbox lubrication model.

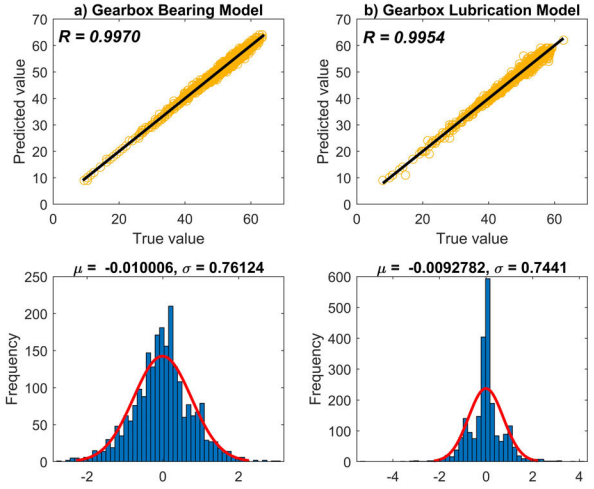

FIGURE 10. The correlation of the predicted and true values and the error distribution histogram of the proposed prediction model in Turbine-A: a) the gearbox bearing model, b) the gearbox lubrication model.

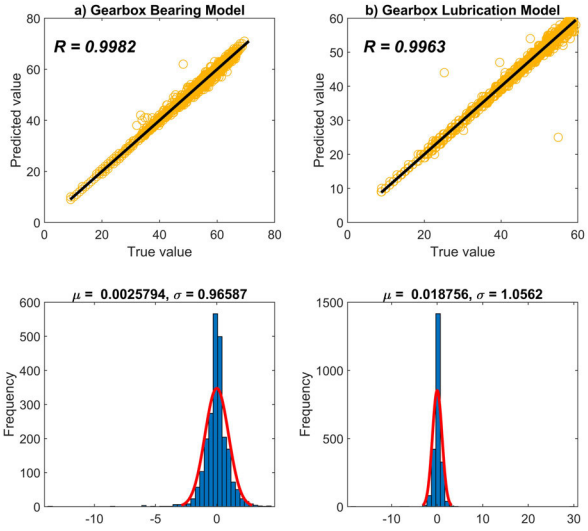

FIGURE 11. The correlation of the predicted and true values and the error distribution histogram of the proposed prediction model in Turbine-B: a) the gearbox bearing model, b) the gearbox lubrication model.

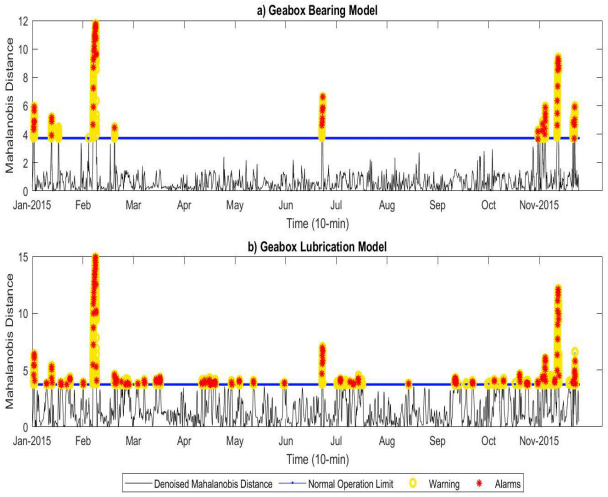

FIGURE 12. The anomaly analysis results during condition monitoring period time for Turbine A- a) gearbox bearing model and b) gearbox lubrication model.

accuracy and performance of the proposed model. Additionally, the error histograms show the error dispersion as well as model probability distribution function.

B. ANOMALY DETECTION ANALYSIS

In this section, the results of fault detection analysis in the wind turbine's gearbox using the gearbox bearing temperature and oil lubrication temperature is presented using real data of wind turbines A and B.

Wind Turbine A: Real SCADA data from January 2013 to December 2014 was used to train the proposed model. The performance of the model was then tested using data from January 2015 to November 2015. During the course of this period, a fault was detected in the pump of gearbox of Turbine A on 11 September 2015. The obtained anomaly detection results based on two target variables (gearbox bearing model and gearbox lubrication model) are highlighted in Figure 12.

According to the results provided in Figure 12, the first alarm in the bearing model occurred on January 1st at 20:40, and the first alarm in lubrication model occurred on February 2nd at 11:20. All the remaining warnings and alarms for the gearbox bearing and lubrication model for different models are illustrated in Table 3. The results show that more warnings and alarms are reported in the lubrication model than the bearing model.

Note that our testing was performed using raw data-sets and no mitigating actions were performed by the operators during the selected time interval, hence the increased number of warnings. Furthermore, in order to better monitor the results of wind turbine A, the monthly warnings and alarms for VMD-GMDH-GA, VMD-GMDH-PSO, and the

TABLE 3. The warnings and alarms information reported in Turbine-A.

Models	Testing period time	Warning		Alarm	
		The bearing model	The lubrication model	The bearing model	The lubrication model
VMD-GMDH-GA	Jan-15	169	278	12	21
VMD-GMDH-PSO		181	210	13	15
Proposed model		157	291	9	20
VMD-GMDH-GA	Feb-15	365	583	27	46
VMD-GMDH-PSO		394	575	29	45
Proposed model		318	571	25	43
VMD-GMDH-GA	Mar-15	93	310	6	22
VMD-GMDH-PSO		76	235	5	18
Proposed model		---	214	---	16
VMD-GMDH-GA	Apr-15	119	302	6	22
VMD-GMDH-PSO		121	331	8	23
Proposed model		---	279	---	19
VMD-GMDH-GA	May-15	55	166	2	11
VMD-GMDH-PSO		63	87	4	6
Proposed model		---	126	---	9
VMD-GMDH-GA	Jun-15	125	277	8	21
VMD-GMDH-PSO		131	234	8	18
Proposed model		87	123	7	9
VMD-GMDH-GA	Jul-15	22	287	1	17
VMD-GMDH-PSO		12	312	1	21
Proposed model		---	265	---	13
VMD-GMDH-GA	Aug-15	7	80	---	5
VMD-GMDH-PSO		---	134	---	9
Proposed model		---	26	---	2
VMD-GMDH-GA	Sep-15	38	214	1	13
VMD-GMDH-PSO		14	241	---	17
Proposed model		---	239	---	16
VMD-GMDH-GA	Oct-15	453	752	29	49
VMD-GMDH-PSO		389	562	27	33
Proposed model		---	158	---	7
VMD-GMDH-GA	Nov-15	219	582	16	41
VMD-GMDH-PSO		143	452	10	30
Proposed model		391	659	29	48

FIGURE 13. Comparison of a) the gearbox bearing model and b) the gearbox lubrication model in terms of warnings.

proposed model are provided in Figures 13 and 14. In these figures, the warnings and alarms rates of different models have been compared with the proposed model for a) the gearbox bearing model and b) the gearbox lubrication model from January 2015 to November 2015. Regarding these results, the maximum of warnings and alarms rates of VMD-GMDH-GA, VMD-GMDH-PSO, and the proposed model are dedicated to October 2015, February 2015, and November 2015 respectively.

Wind Turbine B: Data of wind turbine *B* was selected from January 2014 to December 2014 for training and January 2015 to August 2015 for testing. Note that no fault was reported during the selected monitoring time. The obtained

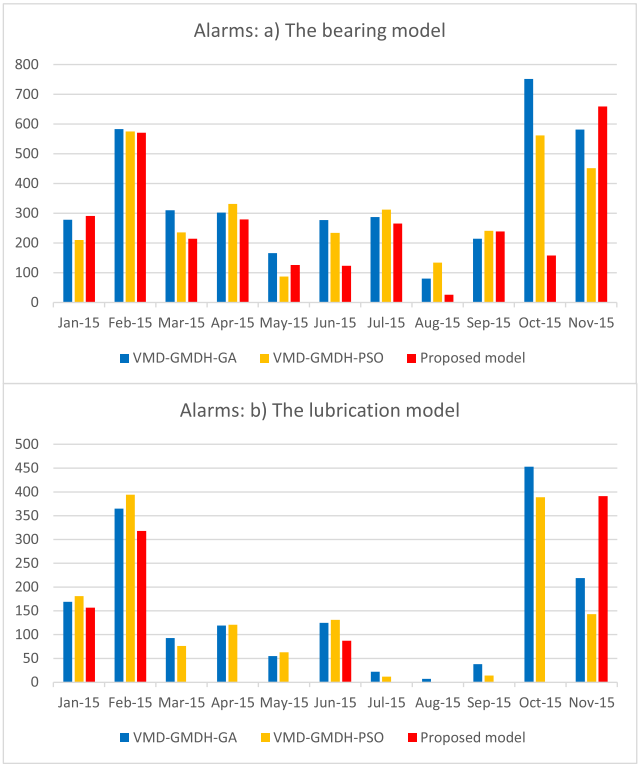

FIGURE 14. Comparison of a) the gearbox bearing model and b) the gearbox lubrication model in terms of alarms details.

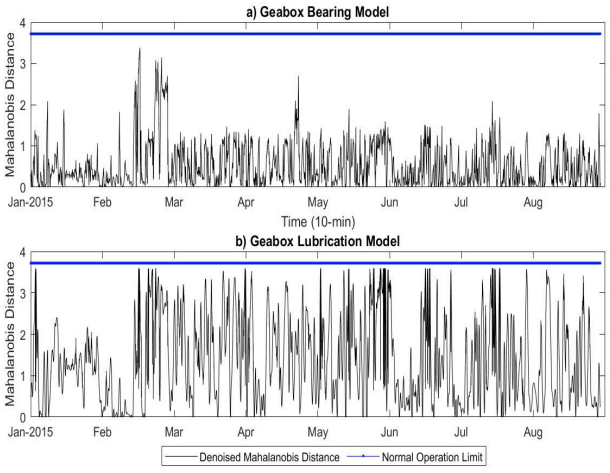

FIGURE 15. The anomaly analysis results during condition monitoring period time for Turbine B- a) gearbox bearing model and b) gearbox lubrication model.

anomaly detection analysis results in this case are illustrated in Figure 15.

The above results indicate that the proposed method did not detect nor report any failure, and are in agreement with the actual condition of wind turbine *B*. This, hence, confirms the prediction accuracy and good performance of the proposed approach.

V. CONCLUSION

This paper proposed a combined intelligent model for the condition monitoring of wind turbine's gearbox. The proposed approach implements a novel hybrid forecasting model combine automatic clustering, variational mode decomposition, GMDH network, and multi-verse optimization algorithm to accurately forecast gearbox bearing and lubrication temperature. In the condition monitoring phase, the predicted and true values have been evaluated by the Mahalanobis distances and wavelet transform de-noising method. The proposed model was implemented to the SCADA data of two on-shore wind turbines. The obtained results showed that the proposed model was able to accurately predict important signals such as bearing and oil lubrication temperatures. Furthermore, in the anomaly detection analysis phase, the proposed model was shown to detect possible anomalies and notify potential operation risks long before failure events, thus preventing unscheduled downtimes, reducing maintenance cost and improving wind turbine's reliability. Our future work will focus on assessing the robustness of the proposed model in the presence of various realistic faulty scenarios.

REFERENCES

[1] D. Y. C. Leung and Y. Yang, "Wind energy development and its environmental impact: A review," *Renew. Sustain. Energy Rev.*, vol. 16, no. 1, pp. 1031–1039, Jan. 2012, doi: 10.1016/j.rser.2011.09.024.

[2] Y. Wang, X. Ma, and P. Qian, "Wind turbine fault detection and identification through PCA-based optimal variable selection," *IEEE Trans. Sustain. Energy*, vol. 9, no. 4, pp. 1627–1635, Oct. 2018, doi: 10.1109/TSTE.2018.2801625.

[3] H. Habibi, I. Howard, and S. Simani, "Reliability improvement of wind turbine power generation using model-based fault detection and fault tolerant control: A review," *Renew. Energy*, vol. 135, pp. 877–896, May 2019, doi: 10.1016/j.renene.2018.12.066.

[4] M. C. Garcia, M. A. Sanz-Bobi, and J. del Pico, "SIMAP: Intelligent system for predictive maintenance," *Comput. Ind.*, vol. 57, no. 6, pp. 552–568, Aug. 2006, doi: 10.1016/j.compind.2006.02.011.

[5] A. Zaher, S. D. J. McArthur, D. G. Infield, and Y. Patel, "Online wind turbine fault detection through automated SCADA data analysis," *Wind Energy*, vol. 12, no. 6, pp. 574–593, Sep. 2009, doi: 10.1002/we.319.

[6] P. Bangalore and L. B. Tjernberg, "An artificial neural network approach for early fault detection of gearbox bearings," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 980–987, Mar. 2015.

[7] M. Schlechtingen, I. F. Santos, and S. Achiche, "Wind turbine condition monitoring based on SCADA data using normal behavior models. Part 1: System description," *Appl. Soft Comput.*, vol. 13, no. 1, pp. 259–270, Jan. 2013, doi: 10.1016/j.asoc.2012.08.033.

[8] L. B. Tjernberg, *Infrastructure Asset Management With Power System Applications*. Boca Raton, FL, USA: CRC Press, 2018.

[9] Z. Hameed, Y. S. Hong, Y. M. Cho, S. H. Ahn, and C. K. Song, "Condition monitoring and fault detection of wind turbines and related algorithms: A review," *Renew. Sustain. Energy Rev.*, vol. 13, no. 1, pp. 1–39, Jan. 2009, doi: 10.1016/j.rser.2007.05.008.

[10] P. Bangalore, S. Letzgus, D. Karlsson, and M. Patriksson, "An artificial neural network-based condition monitoring method for wind turbines, with application to the monitoring of the gearbox," *Wind Energy*, vol. 20, no. 8, pp. 1421–1438, Aug. 2017.

[11] F. Cheng, L. Qu, W. Qiao, C. Wei, and L. Hao, "Fault diagnosis of wind turbine gearboxes based on DFIG stator current envelope analysis," *IEEE Trans. Sustain. Energy*, vol. 10, no. 3, pp. 1044–1053, Jul. 2019.

[12] P. B. Dao, W. J. Staszewski, T. Barszcz, and T. Uhl, "Condition monitoring and fault detection in wind turbines based on cointegration analysis of SCADA data," *Renew. Energy*, vol. 116, pp. 107–122, Feb. 2018, doi: 10.1016/j.renene.2017.06.089.

[13] P. Sun, J. Li, C. Wang, and X. Lei, "A generalized model for wind turbine anomaly identification based on SCADA data," *Appl. Energy*, vol. 168, pp. 550–567, Apr. 2016, doi: 10.1016/j.apenergy.2016.01.133.

[14] H. Sanchez, T. Escobet, V. Puig, and P. Fogh Odgaard, "Fault diagnosis of an advanced wind turbine benchmark using interval-based ARRs and observers," *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3783–3793, Jun. 2015, doi: 10.1109/TIE.2015.2399401.

[15] Y. Zhang, H. Zheng, J. Liu, J. Zhao, and P. Sun, "An anomaly identification model for wind turbine state parameters," *J. Cleaner Prod.*, vol. 195, pp. 1214–1227, Sep. 2018, doi: 10.1016/j.jclepro.2018.05.126.

[16] F. Qu, J. Liu, H. Zhu, and B. Zhou, "Wind turbine fault detection based on expanded linguistic terms and rules using non-singleton fuzzy logic," *Appl. Energy*, vol. 262, Mar. 2020, Art. no. 114469, doi: 10.1016/j.apenergy.2019.114469.

[17] L. Wang, Z. Zhang, H. Long, J. Xu, and R. Liu, "Wind turbine gearbox failure identification with deep neural networks," *IEEE Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1360–1368, Jun. 2017, doi: 10.1109/TII.2016.2607179.

[18] L. M. Halabi, S. Mekhilef, and M. Hossain, "Performance evaluation of hybrid adaptive neuro-fuzzy inference system models for predicting monthly global solar radiation," *Appl. Energy*, vol. 213, pp. 247–261, Mar. 2018, doi: 10.1016/j.apenergy.2018.01.035.

[19] G. Jiang, H. He, J. Yan, and P. Xie, "Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox," *IEEE Trans. Ind. Electron.*, vol. 66, no. 4, pp. 3196–3207, Apr. 2019.

[20] F. Jia, Y. Lei, J. Lin, X. Zhou, and N. Lu, "Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data," *Mech. Syst. Signal Process.*, vols. 72–73, pp. 303–315, May 2016.

[21] H. Zhao, H. Liu, W. Hu, and X. Yan, "Anomaly detection and fault analysis of wind turbine components based on deep learning network," *Renew. Energy*, vol. 127, pp. 825–834, Nov. 2018, doi: 10.1016/j.renene.2018.05.024.

[22] J. Fu, J. Chu, P. Guo, and Z. Chen, "Condition monitoring of wind turbine gearbox bearing based on deep learning model," *IEEE Access*, vol. 7, pp. 57078–57087, 2019, doi: 10.1109/ACCESS.2019.2912621.

[23] P. Santos, L. Villa, A. Reñones, A. Bustillo, and J. Maudes, "An SVM-based solution for fault detection in wind turbines," *Sensors*, vol. 15, no. 3, pp. 5627–5648, Mar. 2015, doi: 10.3390/s150305627.

[24] D. Ma, X. Hu, and H. Zhang, "A hierarchical event detection method based on spectral theory of multidimensional matrix for power system," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 51, no. 4, pp. 2173–2186, Apr. 2021.

[25] Y. Wang, S. Yang, and R. V. Sanchez, "Gearbox fault diagnosis based on a novel hybrid feature reduction method," *IEEE Access*, vol. 6, pp. 75813–75823, 2018, doi: 10.1109/ACCESS.2018.2882801.

[26] Y. Wang, B. Tang, L. Meng, and B. Hou, "Adaptive estimation of instantaneous angular speed for wind turbine planetary gearbox fault detection," *IEEE Access*, vol. 7, pp. 49974–49984, 2019, doi: 10.1109/ACCESS.2019.2908192.

[27] G. D'Amico, F. Petroni, and F. Prattico, "Wind speed prediction for wind farm applications by extreme value theory and copulas," *J. Wind Eng. Ind. Aerodyn.*, vol. 145, pp. 229–236, Oct. 2015, doi: 10.1016/j.jweia.2015.06.018.

[28] Y. He and A. Kusiak, "Performance assessment of wind turbines: Data-derived quantitative metrics," *IEEE Trans. Sustain. Energy*, vol. 9, no. 1, pp. 65–73, Jan. 2018.

[29] 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.

[30] R. K. Pandit and D. Infield, "SCADA based nonparametric models for condition monitoring of a wind turbine," *J. Eng.*, vol. 2019, no. 18, pp. 4723–4727, Jul. 2019, doi: 10.1049/joe.2018.9284.

[31] R. K. Pandit, D. Infield, and A. Kolios, "Gaussian process power curve models incorporating wind turbine operational variables," *Energy Rep.*, vol. 6, pp. 1658–1669, Nov. 2020, doi: 10.1016/j.egyr.2020.06.018.

[32] A. J. R. Reis and A. P. A. da Silva, "Feature extraction via multiresolution analysis for short-term load forecasting," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 189–198, Feb. 2005, doi: 10.1109/TPWRS.2004.840380.

[33] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989, doi: 10.1109/34.192463.

[34] A. J. Conejo, M. A. Plazas, R. Espinola, and A. B. Molina, "Day-ahead electricity price forecasting using the wavelet transform and ARIMA models," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 1035–1042, May 2005, doi: 10.1109/TPWRS.2005.846054.

[35] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014, doi: 10.1109/TSP.2013.2288675.

[36] M. Halkidi and M. Vazirgiannis, "Clustering validity assessment: Finding the optimal partitioning of a data set," in *Proc. IEEE Int. Conf. Data Mining*, Nov. 2001, pp. 187–194, doi: 10.1109/icdm.2001.989517.

[37] S. Das, A. Abraham, and A. Konar, "Automatic hard clustering using improved differential evolution algorithm," *Stud. Comput. Intell.*, vol. 178, no. 1, pp. 137–174, 2009, doi: 10.1007/978-3-540-93964-1_4.

[38] J. C. Dunn, "Well-separated clusters and optimal fuzzy partitions," *J. Cybern.*, vol. 4, no. 1, pp. 95–104, Jan. 1974.

[39] T. Calinski and J. Harabasz, "A dendrite method for cluster analysis," *Commun. Statist. Theory Methods*, vol. 3, no. 1, pp. 1–27, 1974, doi: 10.1080/03610927408827101.

[40] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-1, no. 2, pp. 224–227, Apr. 1979, doi: 10.1109/TPAMI.1979.4766909.

[41] M. K. Pakhira, S. Bandyopadhyay, and U. Maulik, "Validity index for crisp and fuzzy clusters," *Pattern Recognit.*, vol. 37, no. 3, pp. 487–501, Mar. 2004, doi: 10.1016/j.patcog.2003.06.005.

[42] C.-H. Chou, M.-C. Su, and E. Lai, "A new cluster validity measure and its application to image compression," *Pattern Anal. Appl.*, vol. 7, no. 2, pp. 205–220, Jul. 2004, doi: 10.1007/s10044-004-0218-1.

[43] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," Erciyes Univ., Kayseri, Turkey, Tech. Rep. TR06, 2005, p. 10, doi: citeulike-article-id:6592152.

[44] A. G. Ivakhnenko, "Polynomial theory of complex systems," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-1, no. 4, pp. 364–378, Oct. 1971, doi: 10.1109/TSMC.1971.4308320.

[45] G. C. Onwubolu, *Hybrid Self-Organizing Modeling Systems*. Springer, 2009.

[46] S. Atashrouz, G. Pazuki, and Y. Alimoradi, "Estimation of the viscosity of nine nanofluids using a hybrid GMDH-type neural network system," *Fluid Phase Equilibria*, vol. 372, pp. 43–48, Jun. 2014, doi: 10.1016/j.fluid.2014.03.031.

[47] G. Pazuki and S. S. Kakhki, "A hybrid GMDH neural network to investigate partition coefficients of penicillin g acylase in polymer–salt aqueous two-phase systems," *J. Mol. Liquids*, vol. 188, pp. 131–135, Dec. 2013, doi: 10.1016/j.molliq.2013.10.001.

[48] H. Ghanadzadeh, M. Ganji, and S. Fallahi, "Mathematical model of liquid–liquid equilibrium for a ternary system using the GMDH-type neural network and genetic algorithm," *Appl. Math. Model.*, vol. 36, no. 9, pp. 4096–4105, Sep. 2012, doi: 10.1016/j.apm.2011.11.039.

[49] S. Mirjalili, S. M. Mirjalili, and A. Hatamlou, "Multi-verse optimizer: A nature-inspired algorithm for global optimization," *Neural Comput. Appl.*, vol. 27, no. 2, pp. 495–513, Feb. 2016, doi: 10.1007/s00521-015-1870-7.

[50] A. Fathy and H. Rezk, "Multi-verse optimizer for identifying the optimal parameters of PEMFC model," *Energy*, vol. 143, pp. 634–644, Jan. 2018, doi: 10.1016/j.energy.2017.11.014.

[51] Y. Cui, P. Bangalore, and L. B. Tjernberg, "An anomaly detection approach using wavelet transform and artificial neural networks for condition monitoring of wind turbines' gearboxes," in *Proc. Power Syst. Comput. Conf. (PSCC)*, Jun. 2018, pp. 1–7, doi: 10.23919/PSCC.2018.8442916.

[52] Q. Huang, Y. Cui, L. B. Tjernberg, and P. Bangalore, "Wind turbine health assessment framework based on power analysis using machine learning method," in *Proc. IEEE PES Innov. Smart Grid Technol. Eur. (ISGT-Eur.)*, Sep. 2019, pp. 1–5, doi: 10.1109/ISGTEurope.2019.8905495.

AZIM HEYDARI (Student Member, IEEE) was born in Yazd, Iran, in 1989. He is currently pursuing the Ph.D. degree in energy and environment from the Sapienza University of Rome, Italy. In 2019 for six months, he has worked as a Ph.D. Visiting Researcher with the Division of Electric Power and Energy Systems, KTH Royal Institute of Technology, Stockholm, Sweden. He is currently working with the Department of Astronautics, Electrical and Energetic Engineering, Sapienza University of Rome. His research interests include optimization, forecasting, hybrid renewable energy systems, energy management, machine learning, and energy efficiency. He has published several articles and books in reputable academic journals and conference papers.

DAVIDE ASTIASO GARCIA is currently an Assistant Professor of thermal sciences, energy technology, and building physics with the Sapienza University of Rome and the General Secretary of the Italian Wind Energy Association (ANEV). He is also involved as a coordinator of research team in two H2020 projects, such as Geographical Islands FlexibiliTy (GIFT) and Operating a Network of Integrated Observatory Systems in the Mediterranean Sea (ODYSSEA), one EEA project, such as Youth Employment Network for Energy Sustainability in Islands (YENESIS), and one Erasums+project, such as Maintenance Simulator for the Sustainability of European Wind Farms (SIMULWIND). He is contracted by the European Commission Research Executive Agency (REA) as an expert for the evaluation of the project proposal under H2020 calls "FET OPEN (Future Emerging Technologies)." He is an Editor of the *Journal of Energy Research and Reviews* and a Guest Editor of the *Applied Science* and *Frontiers in Energy Research*.

AFEF FEKIH (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the National Engineering School of Tunis, Tunisia, in 1995, 1998, and 2002, respectively. She is currently a Full Professor with the Department of Electrical and Computer Engineering and the Chevron/BORSF Professor in engineering with the University of Louisiana at Lafayette. She has authored or coauthored more than 200 publications in international journals, chapters, and conference proceedings. Her main research interests include control theory and applications, including nonlinear and robust control, optimal control, fault tolerant control with applications to power systems, wind turbines, unmanned vehicles, and automotive engines. She is a member of the Editorial Board of the IEEE Conference on Control Technology and Applications, the IEEE TC on Education, and IFAC TC on Power and Energy Systems.

FARSHID KEYNIA received the Ph.D. degree in power engineering from Semnan University, Iran. He is currently an Associate Professor with the Department of Energy, Institute of Science and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman, Iran. His research interests include energy management, reliability, optimization algorithms, and neural networks applications.

LINA BERTLING TJERNBERG (Senior Member, IEEE) was a Visiting Professor with Stanford University, in 2014, a Postdoctoral Researcher with the University of Toronto, from 2002 to 2003, and a Visiting Researcher with the University of Saskatchewan, in 2000. She is currently a Professor in power grid technology and the Director of the Energy Platform, KTH Royal Institute of Technology, Stockholm, Sweden. She has published over 100 articles and an author of CRC Press with a book on *Infrastructure Asset Management with Power System Applications* (2018). Her research aims to develop models for electric power solutions for the future sustainable energy systems. Her research interests include applied reliability theory and maintenance management. She was a member of the Swedish Government Coordination Council for Smart Grids, from 2012 to 2014. She has served as an expert for the EU commission within Energy, ICT, and Security. She is a Distinguished Lecturer of IEEE Power and Energy Society. She is the Past Chair of the Swedish PE/PEL Chapter. She has served on the Governing Board of IEEE Power and Energy Society and an Editor for the IEEE TRANSACTIONS ON SMART GRID. She chaired the first IEEE ISGT Europe Conference, in 2010.

LIVIO DE SANTOLI is currently a Full Professor of energy management with DIAEE and a Vice Rector for the energy policies with the Sapienza University of Rome. He is an expert in various fields of research, such as renewable energy, energy planning, energy economy, energy efficiency, power to gas, hybrid systems, distributed generation, applied acoustic, air quality, and micro and smart grids. He has more than 200 publications in scientific journals and 12 university books.