CONDITION MONITORING OF WIND TURBINE DRIVETRAINS FOR FAILURE DIAGNOSIS USING DEEP LEARNING WITH AUTOENCODER

Eduardo J.N. Menezes[1,2,*]

[1]Diretor Técnico, Evoluce Tecnologia em Monitoramento e PHM

[2]Centro de Estudos e Ensaios em Risco e Modelagem Ambiental, Universidade Federal de Pernambuco (CEERMA-UFPE)

*Correspondence: eduardo.menezes@evolucetecnologia.com.br

ABSTRACT

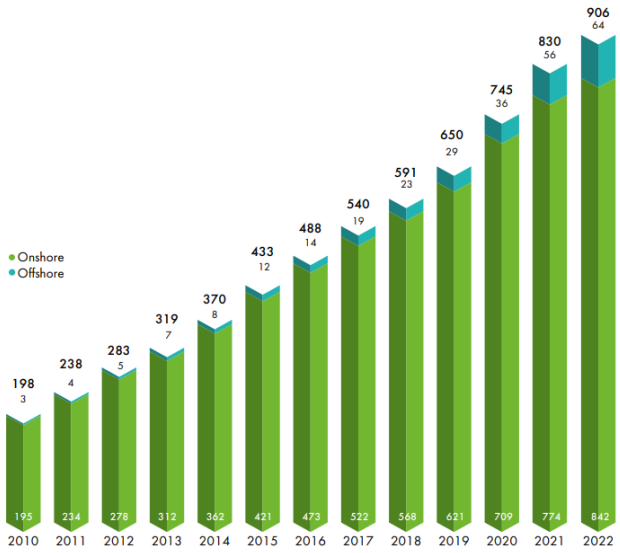
Drivetrain related failures are one of the main concerns for wind turbines (WTs) operation. They are responsible for increased downtimes and associated high O&M costs. Condition monitoring systems based on vibration measurements are the standard in WT applications, but they still rely on vibration analysts and complex frequency-domain analysis to provide diagnostic results. In the present paper, we demonstrate a new diagnosis solution, based on deep learning and direct time-domain feature extraction. The developed tool enables identification of time-series anomalies, indicating the failure warning by surpassing specific thresholds over the reconstruction error of an autoencoder network. Many advantages are demonstrated, including the failure diagnosis long before the traditional FFT and Envelope methods. Effectiveness of the method is proven on a WTs CMS vibration database.

Keywords:
condition monitoring; deep learning; vibration analysis; WT diagnostics; autoencoder.

1. INTRODUCTION

Wind energy has increasingly been incorporated as a major power source in supplying worldwide energy demand, becoming one of the most important and profitable renewable energy sources. Indeed, the world has seen its installed wind power capacity passes from 198 GW in 2010 to 906 GW towards the end of 2022, consisting in a nearly exponential growth of power capacity, as illustrated in Figure 1. On the other hand, the urgent need for energy transition is expected to motivate even more the employ of wind power over many countries of the globe. According to the Intergovernmental Panel on Climate Change (IPCC), organ of the United Nations (UN) for climate changes matters, the limitation of global temperature increase to 2ºC in 2100 is pointed as the superior threshold to ensure tolerable impacts for the ecosystems. This was the limit adopted by the Paris Agreement, in the end of 2015, signed by 192 countries.

Figure 1: Installed wind power capacity approaching an exponential growth trend, onshore and offshore numbers. Source: Adapted by the author from GWEC (2023).



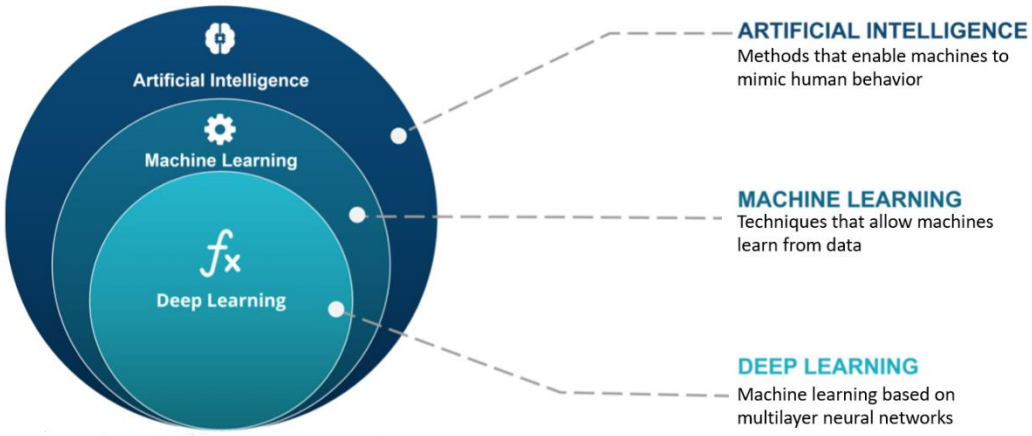
In Brazil, the growth of wind power in the last two decades has proven the potential of including wind energy in the Brazilian energy matrix. The national capacity has increased from the incipient number of 0.22 GW in 2005 to the current 21 GW of wind power production. Brazil is today the 7th country in wind power installed capacity and the projections are of constant increase over the next years (EPBR, 2022). On the other hand, in order to sustain this impulse in wind energy employment, several challenges remain. One of the main reside in the operation and maintenance (O&M) costs, which can attain 30% of the overall wind project cost (NREL, 2020). In fact, wind turbines (WTs) designed for a 20-year lifetime often suffer varied premature failures due to the harsh operational conditions, especially regarding the variable mechanical load (Meng et al., 2019).

Analyzing the failure breakdown, it is clearly demonstrated in many studies that the gearbox and generator failures are the responsible for causing the longest downtime and thus are associated with the highest cost per failure (Dao et al., 2019; Liu & Zhang, 2020). Further, it is noteworthy that bearings are by far the primarily responsible for inducing these drivetrain failures, accounting for, e.g., 76% of the gearbox failures. For the sake of comparison, only 17% of gearbox failures comes from gears and the other 7% from miscellaneous causes (Peeters et al., 2018). It remains evident the importance of avoiding bearing failures, given the high downtimes and costs of drivetrain replacements. The most relevant method to prevent the financial and production losses caused by undesired bearing failures is based on condition monitoring systems (CMS). Although different monitoring variables could be used, the solution largely implemented in the commercial WTs relies on vibration monitoring. Compared to other possible CMSs, such as oil analysis and temperature monitoring, vibration signals are simpler to acquire, provide earlier failure detections and

enable analysis in time, frequency or time-frequency domains (Wymore et al., 2015). Acoustic emissions are comparable to vibration measurement in ability of failure observation, but are harsh to gather and preprocess, especially in noisy environments (Scheeren et al., 2022). However, regardless of vibration CMS being present in the majority of WTs, many difficulties in the diagnosis of drivetrain failures still arise. First, the WT has components working in different frequency ranges, which appear aggregated in the time signal. Second, the essentially random wind loading imposed to the WT structure propagates through the drivetrain, introducing complex vibration patterns in the measured vibration waveform. Third, the analysis and diagnosis of failures using the CMS vibration is still highly dependent on human experts, which uses various signal processing techniques and engineering judgment to indicate failure occurrence.

In order to overcome these issues and improve WT diagnosis, we have been developing several failure detection algorithms, mainly based on machine learning methods. Machine learning (ML) is an artificial intelligence (AI) branch that provides computers with the ability to learn complex relations from data, capturing non-linear and unmodeled dynamics between any number of variables. Thus, ML algorithms are a natural choice for WT diagnosis. Deep learning (DL) is a subfield of machine learning specifically dedicated to algorithms emulating the human brain structure, i.e., neural networks (NNs). Usually, deep learning refers to NNs composed of three or more hidden layers (Deisenroth et al., 2020). The subfields of artificial intelligence are shown in Figure 2.

Figure 2: Artificial Intelligence classification. Source: IBM (2023).



Machine and deep learning have been increasingly studied to execute diagnostics and prognostics of machinery, including WTs, and have shown promising potential. However, very few papers have addressed their application in real turbine data. Therefore, this paper has the following objectives: present one of our diagnostic solutions based on deep learning, using a type of NNs known as autoencoders; apply

the solution in commercial WTs; compare our solution with traditional diagnostic of drivetrain failures, showing it can forecast premature failures in advance to conventional approaches.

2. ML AND DL IN DIAGNOSIS AND PROGNOSIS USING VIBRATION DATA

A large amount of research has been executed in AI for evaluating the health of machines, systems or structures (Berghout & Benbouzid, 2022). This is commonly termed as Prognostics and Health Management (PHM). Various PHM papers rely on laboratory test data which not necessarily corresponds to real field conditions. Nevertheless, they present useful methods that worth consideration and that can be combined and adapted for use in WTs on the field. This paper presents a diagnostic solution which, among other considerations, is also a product from literature survey, and this is why a brief review follows.

PHM techniques for WTs are essentially derived from methods applied to gears and bearing in general rotating machinery (Leite et al., 2018). Older industries such as hydraulic and steam power have a long story of condition monitoring of their components, and therefore they possess very well-established procedures to evaluate equipment's health. Besides, they have smoother operational conditions when compared to WTs, what makes easier the application of general rules, while in the wind industry the solution must be customized. This is another reason for that ML and DL approaches fit so well in WTs analysis, since they are nonlinear by nature. With the increase in computational power and the advent of big data, the last decade has seen an impressive increase in ML and DL research, including using collected vibration data (Do & Söffker, 2021).

One of the first attempts was carried out by Kankar et al. (2011), where the raw vibration data is processed through wavelets. Authors compare the performance of Support Vector Machines (SVMs) with Self-Organizing Maps (SOMs) for bearing defect classification. In the work of Malhi et al. (2011), wavelets are also used, but this time for establishing the remaining useful life (RUL) of bearings using recurrent neural networks (RNNs). A specific application of ML for WTs is first presented by Chen et al. (2013), which presents a neuro-fuzzy method to evaluate pitch faults using the power signal, but without vibration data. The vibration analysis of WTs appears in Zimroz et al. (2014), where the authors propose the use of statistical measures, such as signal RMS, considering the non-stationarity of WT operations. The work aims to improve the recommendations of the german standard VDI 3834 (VDI, 2009). This was the first norm to preview limits for WT vibrations, classifying different zones of attention for vibration RMS levels. Afterwards, in 2015, the classical ISO 10816 has issued its Part 21, containing recommendations for WT vibration analysis. However, these standards are rather prescriptive and general, and their use has limitations. In 2013, Bechhoefer et al. (2013) have made a complete analysis of the Fourier Fast Transform (FFT) as applied to vibration signal monitoring. FFT is largely used in the posterior PHM papers and in WT condition monitoring, as in Verstraete et al. (2017). Other processing techniques were also used for vibration data, such as Hilbert-Huang transform (Soualhi et al., 2015), wavelet packet transform (L. Song et al., 2018) and empirical mode decomposition (Wu, Jiang, et al., 2019). Nevertheless, FFT with envelope analysis is consolidated as the standard tool for WT vibration analysis in the wind industry.

Most recent works have focused on DL methods. In the work of Guo et al. (2017), a health indicator (HI) is constructed based on the training of a RNN using vibration time and time-frequency features, which are ranked by their monotonicity. Authors test their tool in a WT bearing prognosis, achieving a suitable degradation trend, compatible with the field data. Li et al. (2019) have estimated the RUL of bearings with a convolutional neural network (CNN), using as input the short-time Fourier transform (STFT) of the vibration signal, using multi-scale vibration signal. Indeed, the STFT corresponds to the Fourier transform in reduced time window, such that, for every vibration data sample, a 2-D matrix is output. In this context, Wu et al. (2019) propose the construction of a HI based on the bathtub curve, to model the bearing degradation and estimate the RUL. Pecht et al. (2020) employ CNNs with residual blocks of deep NNs, whose input is given by wavelet packet transform. The focus is the classification of bearings and gear failures using vibration data, into 9 different classes. The three classical bearing failures are considered, i.e., inner race, outer race and cage faults. Also, Zhao et al. (2020) have conducted a benchmark analysis of various DL diagnostic methods, where the autoencoder is pointed as one of the promising tools. Autoencoders are also present in the research paper of (Wang et al., 2021), in which the latent space is used for estimating the RUL. The previous papers are focused towards diagnostic and prognostic using vibration signals from laboratory data, mainly from the largely known IEEE-PRONOSTIA platform and the CWRU or MFPT datasets (CWRU, 2013; MFPT, 2016; Nectoux et al., 2012). More extensive reviews regarding DL-based failure detection are presented, e.g., in Zhang et al. (2020) and Rezamand et al. (2020).

Literature review shows that papers focused in providing diagnostic solutions for real WT data is scarce, since there is a necessity of a lot of historical data to implement the developed solutions. This could be executed in a second phase of the CMS implementation, when the operator could use its own historical data to implement estimation of RUL and classification of failure types. First, it is necessary to provide the operator with a tool to identify the WT bearing failures using ML and DL cluster analysis to identify anomalies in the vibration data, and notify the operator to proceed with the condition-based maintenance (CBM). Two papers in this direction are the ones of Ben Ali et al. (2018) and J. Guo et al. (2018). The present paper is aligned with this maintenance philosophy, building a solution applied in real WT CMS vibration data. The developed tool allows the diagnostic of bearing failures, using a deep learning autoencoder and time-signal feature processing, without the need of spectral or envelope analysis, and not depending on direct human intervention.

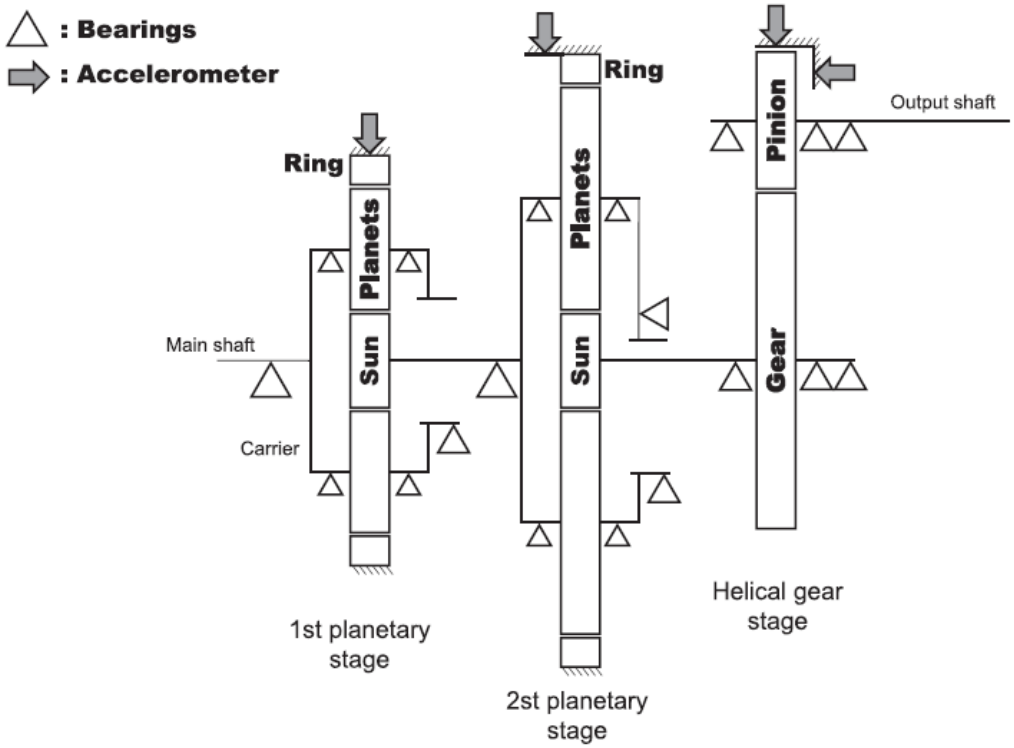
3. METHODOLOGY

3.1 UTILIZED WT CMS DATASET

The applied dataset corresponds to vibration data collected from WT CMSs installed in 18 WTs of 2.5 MW rated capacity over a 5-year period. Data has been provided by Luleå University of Technology (LTU) and is referenced in the work of Strömbergsson et al. (2020). Each gearbox of the WTs is instrumented with 4 accelerometers, mounted in different positions. The gearboxes have three gear stages and two planetary stages, followed by a helical gear stage. The accelerometers are located as follows: one accelerometer at the ring gear of the 1st planetary stage; one accelerometer at the gearbox housing just

beside the ring gear of the 2nd planetary stage; and two accelerometers (radial and axial) at the housing near to the output shaft. Figure 3 illustrates a WT gearbox with the installed accelerometers.

Figure 3: Vibration monitored gearbox with installed accelerometers. Source: (Strömbergsson et al., 2020).

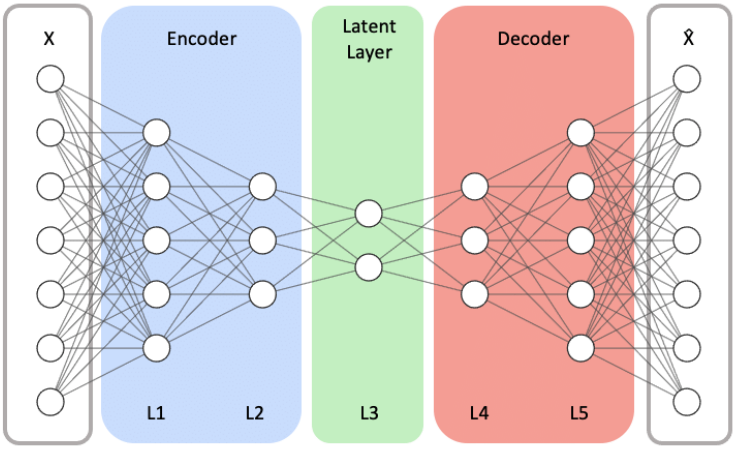


The vibration data has been collected every 24 hours. The accelerometers mounted on the planetary stage had a sample rate of 2.56 kHz and a sample duration of 6.4 s. For the accelerometers mounted on the output shaft, these values were of 12.8 kHz and 1.28 s, where the highest sample frequency is needed due to the higher rotational speed in the output of gearboxes.

3.2 ANOMALY DETECTION USING AUTOENCODERS

Autoencoders (AEs) are a special type of ANNs, whose objective is to reconstruct in the output the same data of the input. In this sense, autoencoders are considered an unsupervised DL method, since the training labels are the input data. The AE is the result of two sequential networks: an encoder, which learns a compressed representation of input data, codifying information to what is called latent space; and a decoder, which obtains the output from the latent space. An example of autoencoder structure is shown in Figure 4.

Figure 4: Autoencoder with sequential encoder and decoder. Source: (Y. Song et al., 2021)



Mathematically, each encoder layer takes the input x and transforms it to a hidden representation h as:

$$h = \phi(W \cdot x + b) \quad (1)$$

where ϕ represents a nonlinear activation function, and W and b are the layer weight and bias, respectively. In this paper, ReLU function will be used as activation for both encoder and decoder layers (see Sec. 3.3.1). The procedure in Eq. 1 generates the latent space and the coded representation. After that, the decoder generates the output x' by passing the data through its several layers, with the same mathematical process, this time acting over the hidden variables:

$$x' = \phi(W' \cdot h + b') \quad (2)$$

where W' and b' are the decoder layer weight and bias. The mean squared error L_{MSE} is normally used as the loss function, as in Eq. 3:

$$L_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(x_i - x_i')^2 \quad (3)$$

where x_i and x_i' are generic samples and N is the total number of samples. In the present paper, the samples consist in vibration time-series collected in each logged measurement of the CMS.

One of the possible uses of AEs is to detect anomalies, based on their reconstruction error (Finke et al., 2021). This process happens when the AEs are trained exclusively with vibration data from healthy machine conditions. In this case, if a non-healthy or anomalous time-series is passed through the network, the AEs will present a large reconstruction error, indicating the anomaly. The procedure has a lot of advantages: it does not require the definition of the defect frequencies for each specific component and failure mode; it does not require feature extraction in the frequency domain; it does not require searching

for specific frequencies or sidebands, which is subjected to human errors. The next section details our AE-based failure diagnosis.

3.3 DEVELOPED TOOL FOR DRIVETRAIN FAILURE DIAGNOSIS

3.3.1 Proposed architecture of the autoencoder

The autoencoder used in our diagnosis tool is composed by 11 layers in the encoder and by the same number of layers in the decoder, since both networks must be symmetric. It is designed to process the vibration signal in time-domain, i.e., employing 1-D sequence data. Indeed, the input of the AE receives sequences of length 4, corresponding to the set of selected time-series features, which are shown in the next section. The AE-network begins with the encoder part, using a Bidirectional long short-term memory (BiLSTM) layer as the first processing step, which allows to capture dependencies between time-steps in both temporal directions, making it useful for complex time-series. BiLSTM is followed by the ReLU activation layer. Afterwards, the AE continues with convolutional 1-D layers, constructed with ReLU activation and dropout layers to prevent overfitting. The latent space, the layer between encoder and decoder, is composed by a single fully-connected layer. After this, a symmetric structure is built for the decoder part, composed by transposed convolutional 1-D and BiLSTMs. At the output, a regression layer based on MSE loss function is utilized. Table 1 shows AE detailed parameters.

Table 1: AE architecture. Source: the Author.

Layer	Architecture
Input	4 time-series features
Encoder	Layer 1: BiLSTM layer (16 hidden units, tanh and sigmoid (state/gate activation functions) Layer 2: ReLU Layer 3: Convolution 1-D layer (16 filters, filter size = 3, stride = 2) Layer 4: ReLU Layer 5: Dropout layer (Probability = 0.2) Layer 6: Convolution 1-D layer (32 filters, filter size = 3, stride = 2) Layer 7: ReLU Layer 8: Dropout layer (Probability = 0.2) Layer 9: Convolution 1-D layer (64 filters, filter size = 3, stride = 1) Layer 10: ReLU Layer 11: Dropout layer (Probability = 0.2)
Latent space	Layer 12: Fully-connected layer (Input size = 64, output size = 1)

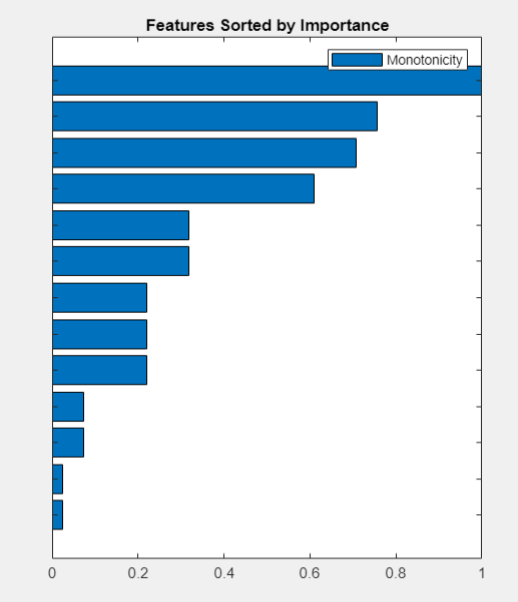
Decoder	Layer 13: Transposed Convolution 1-D layer (64 filters, filter size = 3, stride = 1) Layer 13: ReLU Layer 14: Dropout layer (Probability = 0.2) Layer 15: Transposed Convolution 1-D layer (32 filters, filter size = 3, stride = 2) Layer 16: ReLU Layer 17: Dropout layer (Probability = 0.2) Layer 18: Transposed Convolution 1-D layer (16 filters, filter size = 3, stride = 2) Layer 19: ReLU Layer 20: Dropout layer (Probability = 0.2) Layer 21: BiLSTM layer (16 hidden units, tanh and sigmoid (state/gate activation functions) Layer 22: ReLU Layer 23: Fully-connected`(Input size = 32, output size = 1) Layer 24: Regression layer (MSE loss)

3.3.2 Input data to the AE

The designed AE should use the vibration data directly in time-domain, as collected by the WTs CMS, which was detailed in Sec. 3.1.

Data sequences in the input of the designed AE should correspond to the most meaningful time-series features, in terms of carried information. This allows to represent well the whole time-series, which is composed by 16384 data points, since the CMS has 12.8 kHz of sampling frequency and 1.28-sec time window. For doing that, we execute a procedure of time-features ranking, according to their monotonicity. The ranking makes possible that only a few classical time domain features be enough to run our diagnosis. Therefore, the four top features of higher monotonicity are selected here, in order to better express the degradation in the time domain. The features with evaluated monotonicity were: mean, standard deviation, root mean square, shape factor, kurtosis, skewness, crest factor, impulse factor, peak value, signal-to-noise ratio and total harmonic distortion. The ranking is executed for the diagnosed WT (herein called WT5) and is shown in Figure 6. Upon the features ranking, it is decided that the features input to the AE should be the crest factor, the clearance, the impulse factor and the skewness.

Figure 5: Example of a ranking of time-features (WT5). Source: the Author.



Feature	Monotonicity
Matrix_sigstats/CrestFactor	0.0684
Matrix_sigstats/ClearanceF…	0.0518
Matrix_sigstats/ImpulseFactor	0.0484
Matrix_sigstats/Skewness	0.0417
Matrix_sigstats/Mean	0.0217
Matrix_sigstats/PeakValue	0.0217
Matrix_sigstats/RMS	0.0150
Matrix_sigstats/SNR	0.0150
Matrix_sigstats/Std	0.0150
Matrix_sigstats/ShapeFactor	0.0050
Matrix_sigstats/THD	0.0050
Matrix_sigstats/Kurtosis	0.0017

3.3.3 Drivetrain failure diagnosis using the reconstruction error

Given the chosen input variables, they should be reconstructed when passing through the trained AE network with minimum loss. Further, by executing AE training only on vibration data of healthy WT, the output of AE should indicate anomalous conditions if it is exposed to faulty input data. In a diagnosis tool, we must select reconstruction metrics that provide assurance of the failure detection. For this, we utilize a metric that captures the overall variation between healthy and faulty time-series, as well as a metric that captures the local differences in each sample. The former is the cumulative sum of the mean absolute errors (MAEs) over the last 5 samples (last 5 days) and the latter is the MAE itself within each sample (the daily value). The calculation of both metrics are shown in Eqs. 4 and 5:

$$cMAE = \sum_{i=N-5}^{N} MAE_i \tag{4}$$

$$MAE = \frac{1}{n}\sum_{j=1}^{n}\left\|x_{ji} - x'_{ji}\right\| \tag{5}$$

Where N is the day of analysis, n is the number of features in each sample, and x, x' are the encoder inputs and outputs, respectively. Further, $cMAE$ stands for cumulative MAE. Therefore, the anomaly will be detected from both a global and local time-series analysis. The thresholds of failure condition must be obtained by using the $MAEs$ from the healthy data. As we will see in Sec. 4.2, the extreme MAE and $cMAE$ values found for healthy time-series are employed to this. It is important to observe that the obtained thresholds are specific for each wind farm (WF) and should be calculated by training the AE in the healthy data from its WTs. Indeed, the failure limits will be established through the histograms of $MAEs$ from healthy WTs, without recent failure occurrence (Jonas et al., 2022).

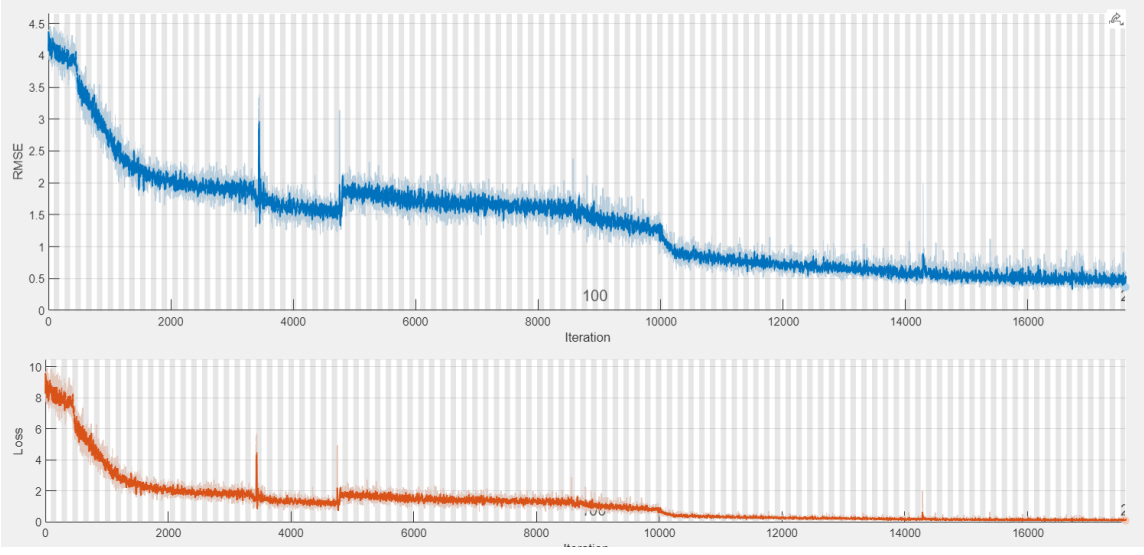
4. RESULTS OF THE DEVELOPED DIAGNOSIS TOOL

It is worth to note the tool can be used with any accelerometer of the CMS, depending on what drivetrain component should be monitored. In the present paper, we will use the accelerometer of the output shaft as a proof-of-concept (PoC). This was the chosen accelerometer because it is related to the failure registered in one of the analyzed WTs. Indeed, it is known from operation logs that the fifth WT (herein called WT5), presented a failure in its output shaft bearing. Operator relates the need for bearing replacement at the measurement number 265 of the used WT5 vibration dataset.

4.1 TRAINING THE AE USING HEALTHY DATA

From the used database, it is known that WT3 and WT4 have not presented any drivetrain failures during the monitoring. Their data are employed for AE training on healthy data and consist in 2837 samples of 16384 data points. The dataset must be split in the training, validation and test datasets, where the latter will be used for comparison with the failure WT5 data. Thus, 70% of WT3 and WT4 data were reserved for train and validation (consisting in 1986 samples), and 30% (851) was reserved for testing. In this way, the AE could be tested in these 851 unseen healthy samples and compared to WT5 failure data. The evolution of training process is illustrated by the RMSE and the AE loss function, as illustrated in Figure 6.

Figure 6: Training metrics of AE expressed in number of iterations (horizontal axis) and RMSE and loss function (vertical axis). Source: the Author.



Note that these error metrics refer to the training phase in the entire healthy dataset and not refer to the failure diagnosis. Backpropagation is used to update the network weights, with Adam optimization algorithm running over the mini-batches. Training occurs in 200 epochs and the complete set of training parameters is described in Table 2.

Table 2: AE training main parameters. Source: the Author.

Autoencoder Training Parameters	
Optimizer	Adam
Initial learn rate	0.001
Learn rate drop factor	0.1
Epsilon	1e-8
Batch size	64
Number of epochs	200

4.2 FAILURE DIAGNOSIS

To test the effectiveness of the designed AE in failure diagnosis and realize our PoC, we submit the data of WT5 output shaft bearing to the trained network. Results for the MAE are shown in Figure 7, compared with healthy data coming from the test set of WT3 and WT4. It can be seen that after the 200th measurement there is the persistence of reconstruction error peaks, with increasing values, indicating a possible anomaly. It is noteworthy that previous error peaks are not persistent or growing, consisting in 1-day higher reconstruction errors, which does not characterize a reconstruction anomaly. A similar situation of peak persistence happens around the 300th measurement. However, the warning regions shown in Figure 7 are the result of pure visual inspection. Failure thresholds must be precise and carefully defined. In order to do this, we use the histograms for MAE and cMAE of healthy data, as shown in Figure 8. We establish the larger values of reconstruction error for the healthy data as failure thresholds. For MAE, we define inferior and superior attention limits, corresponding to the regions where the healthy data histograms present almost no occurrences. Indeed, the immense majority of healthy data locates below MAE < 0.6 and an even larger quantity of data points below MAE < 0.8. Ultimately, beyond the attention limits, the extreme value of healthy data is used as failure threshold at MAE=1.3. For the cMAE, which captures the global effect along 5 consecutive days, we use the same procedure to define attention limits between cMAE=1.4 and cMAE=1.6, while the extreme cMAE=1.9 for healthy data is used as failure threshold. Further, cMAE should be analyzed in conjoint with the MAE to get a complete failure diagnosis. After the analysis to define failure limits in terms of reconstruction error metrics, one can utilize the defined thresholds to run the diagnosis. This is carried out in Figures 9 and 10, where the MAE and cMAE values for the analyzed WT5 drivetrain are presented with the respective limits plotted in horizontal lines.

Figure 7: Comparison between reconstruction errors, showing WT5 potential failure regions. Source: the Author.

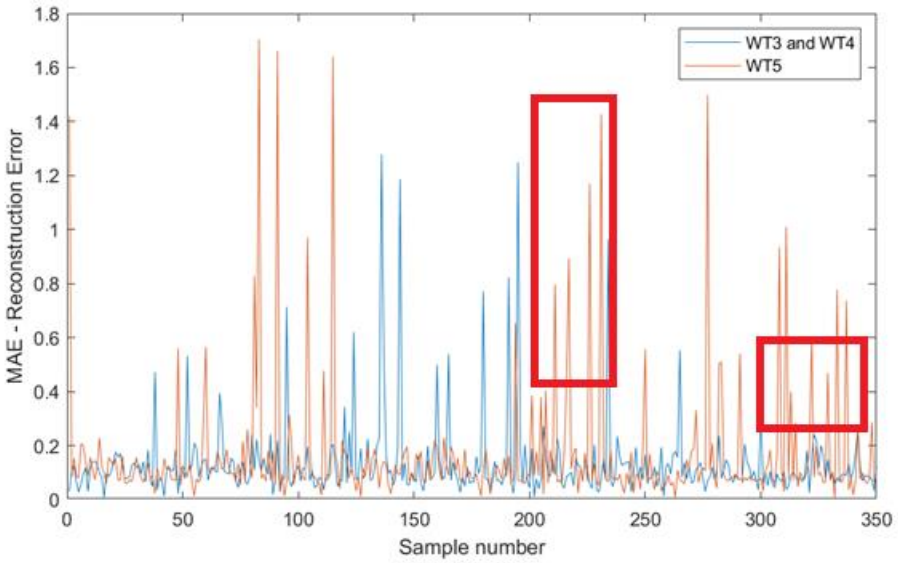
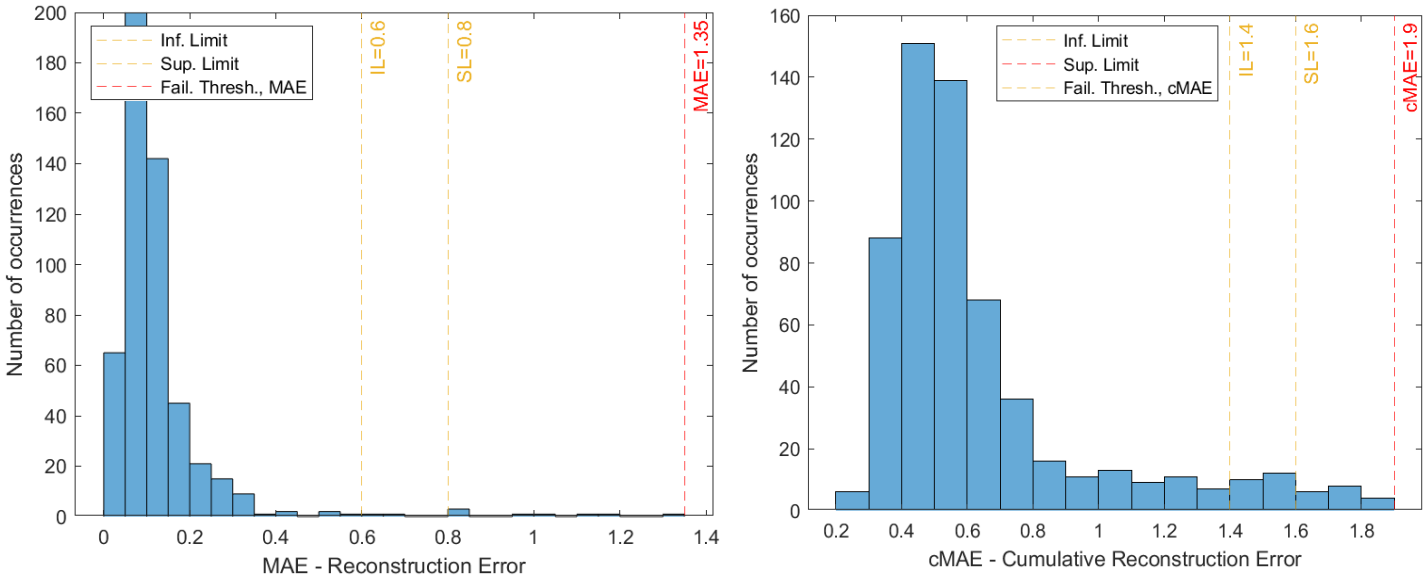


Figure 8: Distribution of occurrences for MAE and cMAE, healthy data. Source: the Author.



In Figure 9, it is possible to note the presence of several measurement regions within the defined attention limits, which consist in potential failure regions. These occur between samples 84-92, 105-116, 195-232 and 309-338. Note that the regions between 84-92 and 105-116 have smaller extension with large discrepancy of subsequent peaks. Regions between measurements 195-232 and 309-338, on the other hand, present a much larger extension and consistency of peaks, with smaller temporal distance between

them and less variation in their amplitude. Additionally, peaks in region 195-232 even present a clear increasing trend, probably associated with the rise of bearing failure. The trend line goes until surpassing the extreme failure threshold (MAE=1.35) at MAE=1.4. Based on this analysis, we could already suppose the measurement 232 as the first failure diagnosis. Nevertheless, we should corroborate this by analyzing the behavior of the cMAE. As shown in Figure 10, a large range of samples within cMAE attention limits are located over the measurements number 195 to 232, confirming the consistent presence of increased reconstruction errors in this region. Further, at measurement 232, the cMAE practically surpasses the extreme value found in the histogram of the right side of Figure 8. After a little intermittency, the cMAE values continues to rise and reaches once again the cMAE threshold at the sample number 274. In addition, observe that from measurements 190-232 and afterwards, there is a trend of increasing in the peak maximum values, attaining cMAE=2.9 in the measurement 381. Combining the results of MAE (representing local error effect) and cMAE (global error effect), we can state confidently that the 232th measurement is the first failure diagnosis, when the operator must cautiously begin the planning of maintenance activities. It is remarkable that in the 232th sample, both failure thresholds are reached and the growing in reconstruction errors have already been set for a long time (since day 190). The day number 232 is the final result of our AE diagnosis tool.

Figure 9: Failure diagnosis chart based on our autoencoder diagnosis tool, MAE analysis. Source: the Author.

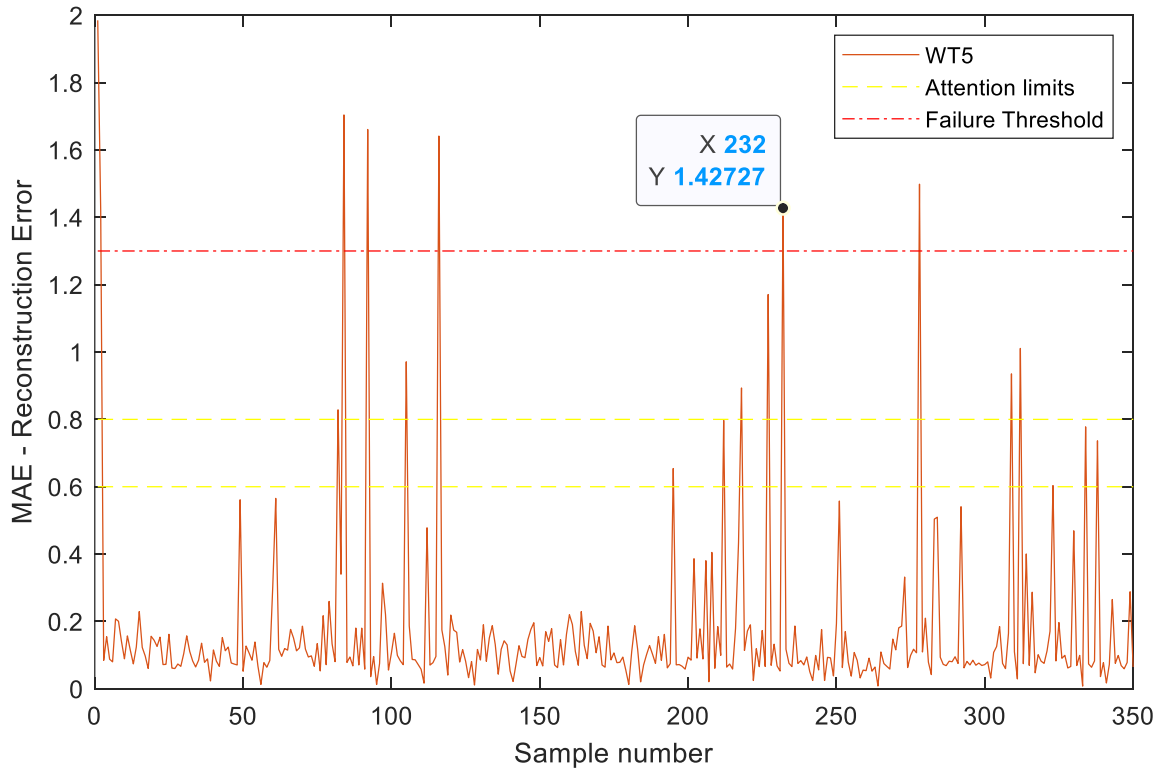
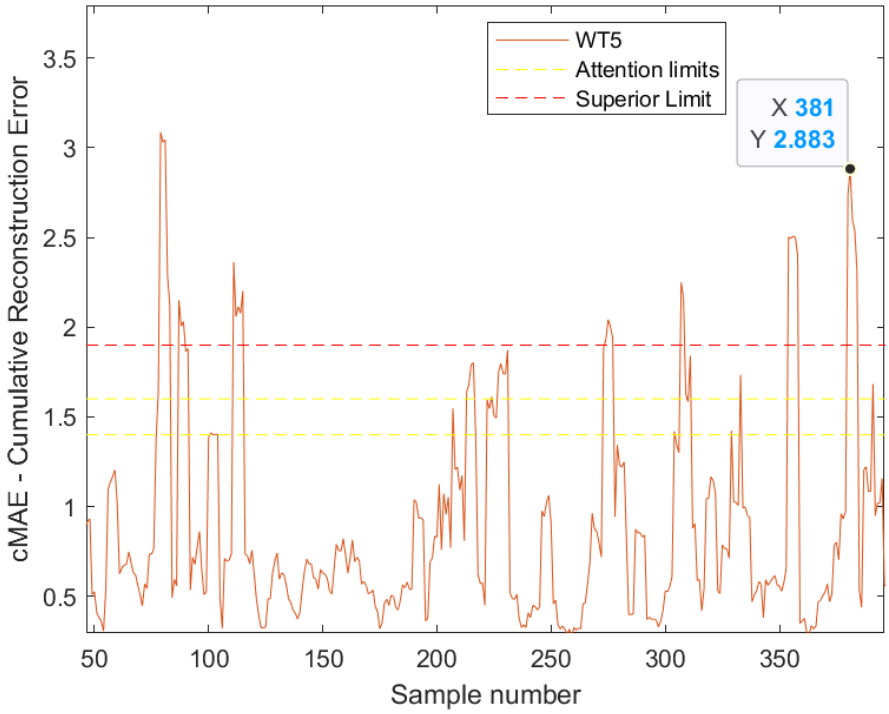


Figure 10: Failure diagnosis chart based on our autoencoder diagnosis tool, cMAE analysis. Source: the Author.



4.3 COMPARISON WITH TRADITIONAL DIAGNOSTIC OF DRIVETRAIN FAILURES

In order to compare the AE-diagnostic with the traditional diagnostic, we analyze WT5 dataset using FFT and envelope analysis. In the present work, the Hilbert transform is executed to obtain the signal envelope. It is well known that the signals of initial bearing degradation are usually encapsulated in high-frequency resonances. The absolute value of Hilbert transform allows to obtain signal demodulation, revealing the bearing defect frequencies, which can be verified by further extracting the FFT of the envelope. More details can be found in, e.g., (Randall, 2014).

The location of the bearing failure in WT5 output shaft was established in the inner race. The calculation for the defect frequency of this specific failure mode has been carried out in the usual way,

$$BPFI = \frac{nf_s}{2}\left\{1 - \frac{d}{D}cos\alpha\right\} \quad (5)$$

Where $BPFI$ is the Ball Pass Frequency Inner race and D, d, α, n and f_s are the pitch diameter, ball diameter, contact angle between the ball and the cage, number of rolling elements, and the rotating speed of bearing (Hz), respectively. For the faulty WT5 bearing, this value was calculated in each rotational speed and monitored carefully. It is important to emphasize that the information of what type of failure is developing is not known a priori. Therefore, in a general case, all the bearing failure frequencies would

need to be monitored (and not only the BPFI), which makes the failure diagnosis even more difficult. This is another advantage of our AE diagnosis tool, which is independent of the failure type.

In Figure 10, we can see the FFT of the envelope fat the beginning of WT5 operation, with the indicated BPFI amplitude. Executing the WT5 monitoring, it is observed that the failure is not evident until approximately the measurement of number 260. The FFT of the envelope for the faulty bearing is shown in Figure 11. As the designed AE indicated the beginning of failure at measurement of number 232, this represents a 28-days advance in drivetrain warning.

Figure 11: FFT of WT5 envelope, measurement number 50. Source: the Author.

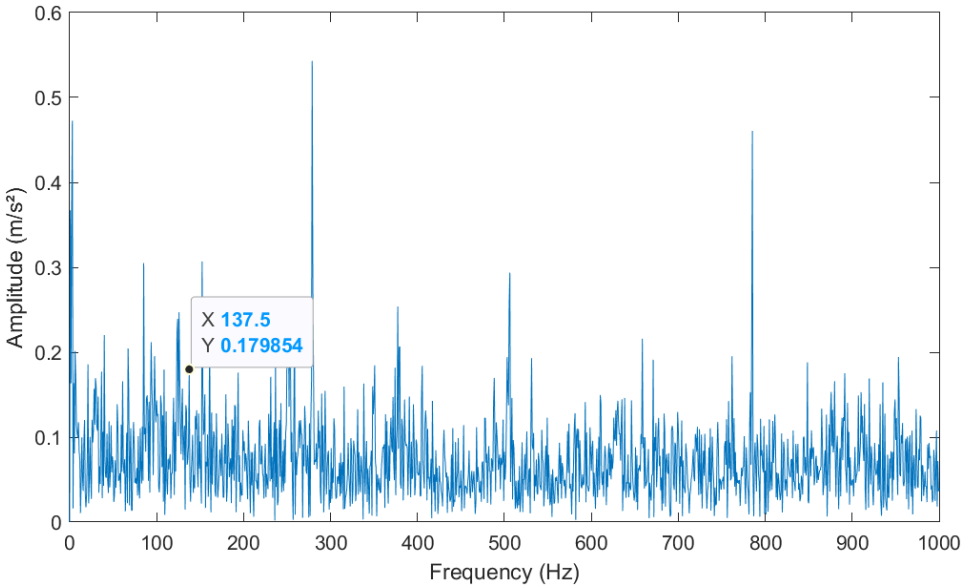
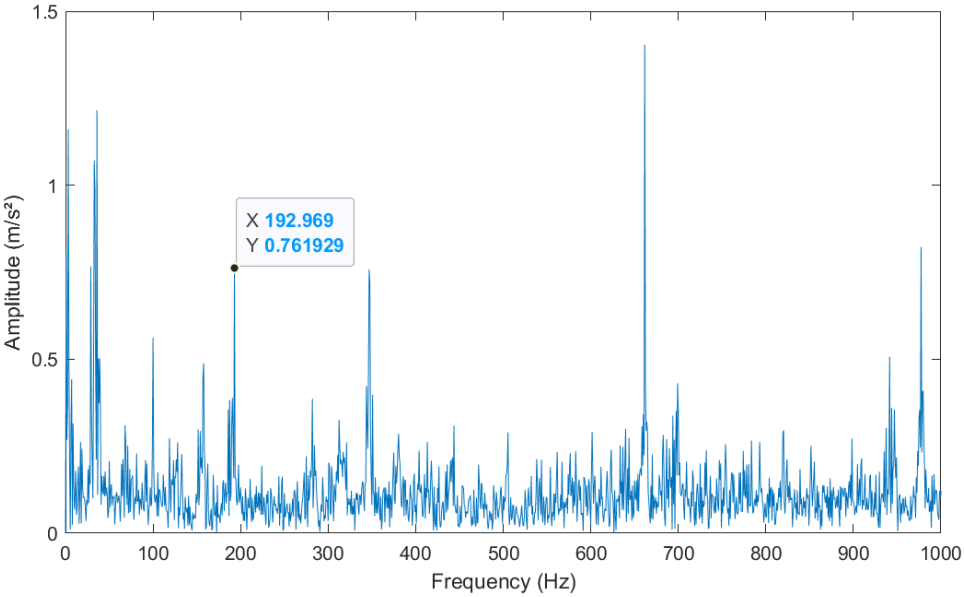


Figure 12: FFT of WT5 envelope, measurement number 260. Source: the Author.



5. CONCLUSIONS

In the present paper, we have presented a diagnostic solution for WT drivetrains considering vibration CMS. Our tool is based on the design of a deep learning AE, composed by 24 layers, which

processes the signal directly in the time domain. This eliminates the need for tracking the specific frequency of drivetrain defects, sometimes hard to obtain and distinguish in the complex vibration environment of a WT. Further, the procedure also issues warnings of possible drivetrain problems long before the traditional vibration analysis. This happens due to the ability of the deep AE to learn the non-linear and intricate relations within the time-series, providing means of identification of what is a healthy and what is an anomaly condition. To sum up, the herein presented tool achieves the following outcomes:

- Identification of failures direct from time-series, not relying on frequency feature extraction and specific frequency defects.
- Independence of traditional FFT-envelope monitoring of each sample.
- Being an unsupervised learning method, the designed AE does not depend on the availability of previous failure data. This allows the use of the method since the start of WF operation.
- Identification of failures much earlier (approximately 28-days before) than the traditional diagnosis methods.
- Possibility of detecting failures in other components or sub-systems, by the re-training of the AE on corresponding data.

Future works comprise the application of the developed tool in other WT databases. Also, the use of other CMSs, such as temperature monitoring, is envisaged, as well as the inclusion of several SCADA data as part of the diagnosis procedures. At last, we plan to continue our work over the prognostic of WT components, estimating the RUL with especially designed DL-networks.

REFERENCES

Bechhoefer, E., Van Hecke, B., & He, D. (2013). Processing for Improved Spectral Analysis.

Ben Ali, J., Saidi, L., Harrath, S., Bechhoefer, E., & Benbouzid, M. (2018). Online automatic diagnosis of wind turbine bearings progressive degradations under real experimental conditions based on unsupervised machine learning. Applied Acoustics, 132, 167–181. https://doi.org/10.1016/j.apacoust.2017.11.021

Berghout, T., & Benbouzid, M. (2022). A Systematic Guide for Predicting Remaining Useful Life with Machine Learning. In Electronics (Switzerland) (Vol. 11, Issue 7). MDPI. https://doi.org/10.3390/electronics11071125

Chen, B., Matthews, P. C., & Tavner, P. J. (2013). Wind turbine pitch faults prognosis using a-priori knowledge-based ANFIS. Expert Systems with Applications, 40(17), 6863–6876. https://doi.org/10.1016/j.eswa.2013.06.018

CWRU. (2013). Case Western Reserve University (CWRU) Bearing Data Center.

Dao, C., Kazemtabrizi, B., & Crabtree, C. (2019). Wind turbine reliability data review and impacts on levelised cost of energy. In Wind Energy (Vol. 22, Issue 12, pp. 1848–1871). John Wiley and Sons Ltd. https://doi.org/10.1002/we.2404

Deisenroth, M. P., Faisal, A. A., & Ong, C. S. (2020). Mathematics for Machine Learning. Cambridge University Press. https://books.google.com.br/books?id=pbONxAEACAAJ

Do, M. H., & Söffker, D. (2021). State-of-the-art in integrated prognostics and health management control for utility-scale wind turbines. Renewable and Sustainable Energy Reviews, 145. https://doi.org/10.1016/j.rser.2021.111102

EPBR. (2022). Eólica chega a 21 GW de capacidade instalada no Brasil. https://epbr.com.br/eolica-chega-a-21-gw-de-capacidade-instalada-no-brasil/

Finke, T., Krämer, M., Morandini, A., Mück, A., & Oleksiyuk, I. (2021). Autoencoders for unsupervised anomaly detection in high energy physics. Journal of High Energy Physics, 2021(6), 161. https://doi.org/10.1007/JHEP06(2021)161

Guo, J., Lu, S., Zhai, C., & He, Q. (2018). Automatic bearing fault diagnosis of permanent magnet synchronous generators in wind turbines subjected to noise interference. Measurement Science and Technology, 29(2). https://doi.org/10.1088/1361-6501/aa92d6

Guo, L., Li, N., Jia, F., Lei, Y., & Lin, J. (2017). A recurrent neural network based health indicator for remaining useful life prediction of bearings. Neurocomputing, 240, 98–109. https://doi.org/10.1016/j.neucom.2017.02.045

Jonas, S., Anagnostos, D., Brodbeck, B., & Meyer, A. (2022). Vibration fault detection in wind turbines based on normal behaviour models without feature engineering.

Kankar, P. K., Sharma, S. C., & Harsha, S. P. (2011). Rolling element bearing fault diagnosis using wavelet transform. Neurocomputing, 74(10), 1638–1645. https://doi.org/10.1016/j.neucom.2011.01.021

Leite, G. de N. P., Araújo, A. M., & Rosas, P. A. C. (2018). Prognostic techniques applied to maintenance of wind turbines: a concise and specific review. In Renewable and Sustainable Energy Reviews (Vol. 81, pp. 1917–1925). Elsevier Ltd. https://doi.org/10.1016/j.rser.2017.06.002

Li, X., Zhang, W., & Ding, Q. (2019). Deep learning-based remaining useful life estimation of bearings using multi-scale feature extraction. Reliability Engineering and System Safety, 182, 208–218. https://doi.org/10.1016/j.ress.2018.11.011

Liu, Z., & Zhang, L. (2020). A review of failure modes, condition monitoring and fault diagnosis methods for large-scale wind turbine bearings. In Measurement: Journal of the International Measurement Confederation (Vol. 149). Elsevier B.V. https://doi.org/10.1016/j.measurement.2019.107002

Malhi, A., Yan, R., & Gao, R. X. (2011). Prognosis of defect propagation based on recurrent neural networks. IEEE Transactions on Instrumentation and Measurement, 60(3), 703–711. https://doi.org/10.1109/TIM.2010.2078296

Meng, H., Lien, F. S., Glinka, G., & Geiger, P. (2019). Study on fatigue life of bend-twist coupling wind turbine blade based on anisotropic beam model and stress-based fatigue analysis method. Composite Structures, 208, 678–701. https://doi.org/10.1016/J.COMPSTRUCT.2018.10.032

MFPT. (2016). Society for Machinery Failure Prevention Technology.

Nectoux, P., Gouriveau, R., Medjaher, K., Ramasso, E., Morello, B., Zerhouni, N., & Varnier, C. (2012). PRONOSTIA: An Experimental Platform for Bearings Accelerated Degradation Tests.

Pecht, M., Zhao, M., Tang, B., & Deng, L. (2020). Multiple wavelet regularized deep residual networks for fault diagnosis. Measurement: Journal of the International Measurement Confederation, 152. https://doi.org/10.1016/j.measurement.2019.107331

Peeters, C., Guillaume, P., & Helsen, J. (2018). Vibration-based bearing fault detection for operations and maintenance cost reduction in wind energy. Renewable Energy, 116, 74–87. https://doi.org/10.1016/j.renene.2017.01.056

Randall, E. R. B. (2014). Machine Diagnostics using Advanced Signal Processing.

Rezamand, M., Kordestani, M., Carriveau, R., Ting, D. S. K., Orchard, M. E., & Saif, M. (2020). Critical Wind Turbine Components Prognostics: A Comprehensive Review. In IEEE Transactions on Instrumentation and Measurement (Vol. 69, Issue 12, pp. 9306–9328). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/TIM.2020.3030165

Scheeren, B., Kaminski, M. L., & Pahlavan, L. (2022). Evaluation of Ultrasonic Stress Wave Transmission in Cylindrical Roller Bearings for Acoustic Emission Condition Monitoring. Sensors, 22(4). https://doi.org/10.3390/s22041500

Song, L., Wang, H., & Chen, P. (2018). Vibration-Based Intelligent Fault Diagnosis for Roller Bearings in Low-Speed Rotating Machinery. IEEE Transactions on Instrumentation and Measurement, 67(8), 1887–1899. https://doi.org/10.1109/TIM.2018.2806984

Song, Y., Hyun, S., & Cheong, Y. G. (2021). Analysis of autoencoders for network intrusion detection†. Sensors, 21(13). https://doi.org/10.3390/s21134294

Soualhi, A., Medjaher, K., & Zerhouni, N. (2015). Bearing health monitoring based on hilbert-huang transform, support vector machine, and regression. IEEE Transactions on Instrumentation and Measurement, 64(1), 52–62. https://doi.org/10.1109/TIM.2014.2330494

Strömbergsson, D., Marklund, P., Berglund, K., & Larsson, P. E. (2020). Bearing monitoring in the wind turbine drivetrain: A comparative study of the FFT and wavelet transforms. Wind Energy, 23(6), 1381–1393. https://doi.org/10.1002/we.2491

VDI. (2009). VDI 3834 Blatt 1 - Wind turbines with gearbox.

Verstraete, D., Ferrada, A., Droguett, E. L., Meruane, V., & Modarres, M. (2017). Deep learning enabled fault diagnosis using time-frequency image analysis of rolling element bearings. Shock and Vibration, 2017. https://doi.org/10.1155/2017/5067651

Wang, C., Jiang, W., Yang, X., & Zhang, S. (2021). Rul prediction of rolling bearings based on a dcae and cnn. Applied Sciences (Switzerland), 11(23). https://doi.org/10.3390/app112311516

Wu, C., Feng, F., Wu, S., Jiang, P., & Wang, J. (2019). A method for constructing rolling bearing lifetime health indicator based on multi-scale convolutional neural networks. Journal of the Brazilian Society of Mechanical Sciences and Engineering, 41(11). https://doi.org/10.1007/s40430-019-2010-6

Wu, C., Jiang, P., Ding, C., Feng, F., & Chen, T. (2019). Intelligent fault diagnosis of rotating machinery based on one-dimensional convolutional neural network. Computers in Industry, 108, 53–61. https://doi.org/10.1016/j.compind.2018.12.001

Wymore, M. L., Van Dam, J. E., Ceylan, H., & Qiao, D. (2015). A survey of health monitoring systems for wind turbines. In Renewable and Sustainable Energy Reviews (Vol. 52, pp. 976–990). Elsevier Ltd. https://doi.org/10.1016/j.rser.2015.07.110

Zhang, S., Zhang, S., Wang, B., & Habetler, T. G. (2020). Deep Learning Algorithms for Bearing Fault Diagnosticsx - A Comprehensive Review. In IEEE Access (Vol. 8, pp. 29857–29881). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2020.2972859

Zhao, Z., Li, T., Wu, J., Sun, C., Wang, S., Yan, R., & Chen, X. (2020). Deep learning algorithms for rotating machinery intelligent diagnosis: An open source benchmark study. ISA Transactions, 107, 224–255. https://doi.org/10.1016/j.isatra.2020.08.010

Zimroz, R., Bartelmus, W., Barszcz, T., & Urbanek, J. (2014). Diagnostics of bearings in presence of strong operating conditions non-stationarity - A procedure of load-dependent features processing with application to wind turbine bearings. Mechanical Systems and Signal Processing, 46(1), 16–27. https://doi.org/10.1016/j.ymssp.2013.09.010