

# A Hybrid Model for Wind Turbine Main Bearing Fatigue with Uncertainty in Grease Observations

Yigit A. Yucesan<sup>1</sup> and Felipe A. C. Viana<sup>2</sup>

<sup>1,2</sup> *Department of Mechanical and Aerospace Engineering, University of Central Florida, Orlando, FL, 32816, USA*  
yucesan@knights.ucf.edu  
viana@ucf.edu

## ABSTRACT

Available historical field data shows that wind turbine main bearing failure can lead to major operation and maintenance costs due to unscheduled downtime. For legacy turbines, fatigue is one of the major failure modes and, to a degree, can be partially modeled with physics-based formulations. Unfortunately, existing bearing fatigue models can potentially be inaccurate due to lack of understanding of the lubricant degradation. One way to enhance these models is to track the grease damage along with the bearing fatigue damage. However, the need of grease degradation data can become an impediment for such strategy. In this paper, we will demonstrate that it is possible to calibrate grease degradation models with cost-efficient periodic visual inspections. Knowing that such inspections introduce observation uncertainty to the model, we will use a hybrid physics-informed deep neural networks to quantify such uncertainties within our models. We built a hybrid model that fuses the physics-based understanding of the bearing fatigue failure with the ability of data-driven layers to compensate the missing physics, with respect to the grease degradation. The proposed hybrid model is also capable of decoding uncertain visual grease inspections with a custom designed classifier. We illustrate the merits of the model with the support of case studies, where we test inspection with different levels of conservatism to train the model and compare the predictions of these models on an artificial wind park. Results from the case studies indicate the successful prognostic performance of the trained with limited and noisy observations. While grease damage is predicted with 0.3% root mean square error as a result of baseline inspection campaign, bearing life prediction is conservatively off only by months for aggressive turbines that have 10 years of life.

## 1. INTRODUCTION

Wind turbine main bearing is a major focus in terms of reliability, since unexpected failures result in costly maintenance and undesired downtime. Hornemann and Crowther (2013) draw attention to the multiple failure modes that main bearings are subjected to. Common failure modes described for wind turbine bearings include, but are not limited to, micro-pitting, white etching crack, electrical erosion, and contact fatigue or spalling. Inherent manufacturing issues tend to dominate early life failures, in addition to issues related to extreme loads and environmental conditions, and maintenance practices. As the fleet is aged, these issues are often mitigated and the dominant failure mode becomes the fatigue (Sethuraman et al., 2015).

Fatigue life modeling of bearings is a topic that is widely researched. Watanabe and Uchida (2015) built a model to predict wind turbine rear bearing fatigue using standard bearing life calculations. Their model proved good agreement with the actual failures observed in a wind site located in Japan. Whereas the collected actual field failure data for a specific turbine is 12.7 years, the model predicted 12 years. The authors also provided ways to utilize their model to manage life extension operations (i.e. curtailment). We should also note that the referred paper consider wind turbines with four-point mounting setting, where two main bearings exist (front and rear), and focus on rear main bearings. For comparison, while in four-point mounting setting two bearings share the load, in three-point mounting setting single main bearing carry the entire incoming load. Walker and Coble (2018) proposed a method that is the combination of adaptive sampling and order tracking approach to examine vibration data for main bearing anomaly detection. In the provided case study, the model was able to detect the bearing fault and the post-mortem examinations have confirmed the failure. Yucesan and Viana (2020) used a physics-informed neural networks model encapsulated into a recurrent neural network to estimate main bearing fatigue life. The authors used physics-based formulations for fatigue modeling, while the grease degradation is modeled by a data-driven node. The case study results prove that the data-

---

Yigit A. Yucesan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

driven portion of the model can capture the grease degradation behavior, by only providing small number of grease damage observations.

Researchers have also studied grease, as lubrication efficiency is related to useful life of bearings. Zhu et al. (2013) performed lubricant state identification by merging several different techniques. The authors built a model to predict remaining useful life of lubricant that utilizes viscosity of the lubricant and dielectric constant sensor output with particle filtering technique. Performance of their approach is evaluated with experiments, which concluded that single observation on dielectric constant output provides the most accurate estimation on lubricant life. Yucesan and Viana (2019a) used cumulative damage model accounting for grease degradation to assess the impact of asset-specific regreasing policies. They built a physics-based fatigue life model and tested turbine-specific regreasing policies across different farms. The results from the case studies indicate that significant life extension can be achieved through turbine-specific regreasing.

With that said, we can see how building remaining useful life models for main bearings is a challenging task. Even though bearing fatigue is relatively well understood, limited assessment and monitoring bearing damage during operation and poor understanding of grease degradation mechanism makes it difficult to build robust models that have accurate predictions.

**In this contribution, we will focus on the issue of reducing the discrepancy of the bearing fatigue model, due to lack of knowledge about grease degradation, with observations coming from visual grease inspection.** We propose using a hybrid physics-informed neural network model that fuses knowledge of the bearing fatigue with a data-driven kernels to compensate unknown grease degradation mechanism. The resulting deep neural network handles uncertainty in grease visual inspection by mapping inspection readings into damage. We build an artificial wind farm damage history for both bearing fatigue and grease using physics-models and manufacturer catalogs to use as a ground true of our case study. With the help of a numerical experiment, we will address the following fundamental question: how accurate is the resulting physics-informed neural networks model built with grease visual inspection?

Physics-informed neural network modeling has received growing attention of researchers over the past few years. For example, Raissi (2018) proposed using two deep learning networks in order to approximate a partial differential equation in fluid mechanics. While one model is used as a prior to the system solution, the other one is used as fine tuning to steady state solution. Bergs et al. (2018) also studied and introduced methods to fuse data-driven models with theoretical models to benefit both approaches for enhanced predictive capabilities. On the other hand, there is work on building hybrid mod-

els that directly code reduced order physics-informed models within neural networks (Nascimento & Viana, 2019; Dourado & Viana, 2020). This is the focus here. We will extend our previous work (Yucesan & Viana, 2020) so that the hybrid model is robust to uncertainty in visual inspection.

The remaining of the paper is organized as follows. Section 2 gives an overview on challenges of modelling main bearing fatigue damage accumulation under uncertain grease inspections. Section 3 elaborates the physics-informed neural networks model we propose as the solution to the problem. Section 4 describes the case study with regards to the wind farm and inspection campaigns. Section 5 presents and discusses the numerical results. Finally, section 6 concludes the paper by summarizing significant remarks, and providing insight on potential future studies. There is one appendix at the end of the paper, discussing neural networks weight initialization, multi-layer perceptrons, benchmark study of proposed hybrid method against pure data-driven models, baseline grease degradation data, and input data preprocessing.

## 2. MODELING MAIN BEARING FATIGUE AND CHALLENGES ASSOCIATED WITH GREASE VISUAL INSPECTION

### 2.1. Main Bearing Fatigue Damage Accumulation

In this paper, we will use the same modeling strategy presented in Yucesan and Viana (2020). While in that paper, we describe the bearing fatigue cumulative damage model in great detail, here we will only highlight the main features. For spherical roller bearings operating at different load levels and rotational speeds, fatigue damage,  $a_{BRG}$ , is governed by (SKF-contributors, 2007):

$$\begin{aligned} \frac{da_{BRG}}{dt} &= \frac{1}{c_1 c_2(t)} \left( \frac{P(t)}{C} \right)^{\frac{10}{3}}, \\ P(t) &= f_1(V_W(t)), \\ c_2(t) &= f_2(P(t), \eta_c(t), \nu(t)), \\ \eta_c(t) &= f_3(\nu(t), a_{GRS}(t)), \text{ and} \\ \nu(t) &= f_4(T_{BRG}(t), a_{GRS}(t)), \end{aligned} \quad (1)$$

where  $c_1$  is a reliability level factor (see Tab. 1);  $c_2$  is an adjustment factor;  $P$  is the equivalent dynamic bearing load;  $C$  is the design load rating;  $\eta_c$  is the grease contamination factor;  $\nu$  is the viscosity;  $V_W$  is the wind speed;  $T_{BRG}$  is the bearing temperature;  $a_{GRS}$  is an indicator of grease degradation; and  $f_{1..4}(\cdot)$  are functions defining the models for different components of the bearing damage.

In this paper, we study a 1.5 MW wind turbine with 80 meters hub height, equipped with a main bearing in the three-point mounting configuration -- further details given by GE-contributors (2009) and SKF-contributors (2007).  $P$  and  $c_2$  vary over time due to wind speed and bearing temperature,

Table 1.  $c_1$  bearing fatigue life adjustment factor (SKF-contributors, 2007).

| Reliability level (%) | 90   | 95   | 97   | 98   | 99   |
|-----------------------|------|------|------|------|------|
| $c_1$                 | 1.00 | 0.62 | 0.44 | 0.33 | 0.21 |

as well as grease condition, which strongly contributes to bearing damage accumulation. Figure 1a shows how bearing loads varies with wind speed (i.e.,  $f_1$  in Eq. 1) -- results obtained through high-fidelity multi-body dynamics analysis (Sethuraman et al., 2015). Figure 1b illustrates the input-output relationship for  $\eta_c$ ,  $\nu$ , and  $c_2$  (i.e.,  $f_{2...4}$  in Eq. 1).

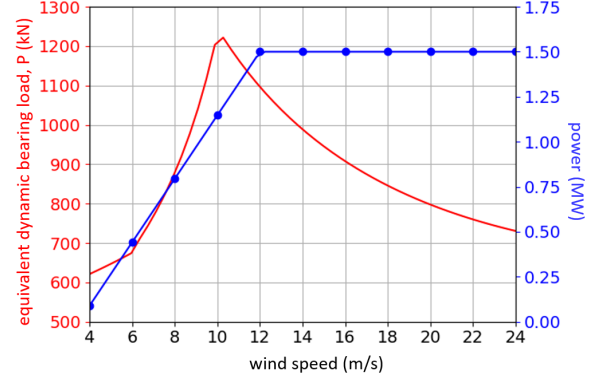
## 2.2. Grease Visual Inspection

In practice, there are multiple ways a wind park operator can conduct an inspection campaign in order to assess the state of the grease. Potentially, the most accurate methodology is to extract a sample from the machine, and have a laboratory conducting detailed analysis. These tests could provide data about the state of the lubricant in terms of viscosity and several other indexes of grease degradation and contamination. Unfortunately, laboratory tests are usually expensive and time consuming (not to mention that results can be biased by procedure used to collect the grease samples).

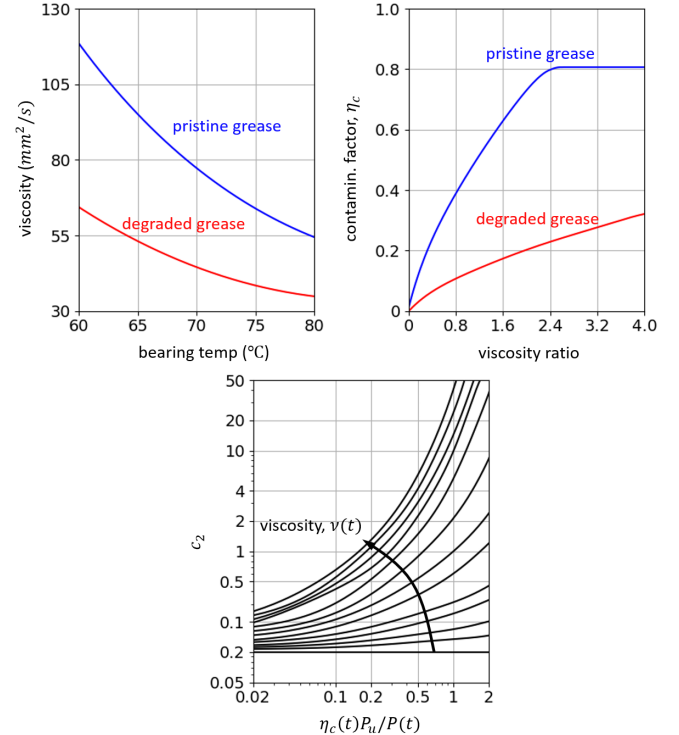
Alternatively, operators can opt for visual inspection performed by trained technicians as an affordable and fast approach to monitor grease degradation. Clearly, the cost and speed advantages of visual inspection come at the cost of the large uncertainties associated with it. Visual inspection of a lubricant is essentially the judgement of the technician on the grease state, based on the visual indicators of the lubricant, mainly the color and the contamination that is visually available to the naked eye. In addition, such assessment is rather prone to human error (or intentional or unintentional conservatism in visual assessment), which can impose challenges for later use of the data (specially in modeling). Various other factors may also affect the technician ranking accuracy, such as poor lighting and the field of view besides human subjectivity. Figure 2 illustrates a potential ranking system for the current quality of the lubricant based on the visual hints. Note that the ranking system is discrete (grease quality ranked from 1 to 5), as opposed to the detailed quantitative results that could come out of a laboratory test.

In summary, from a modeling perspective, these are the major challenges imposed by grease visual inspection:

- visual grease inspection is subjected to large variability due to technician reading;
- reinforcing consistency in technician can be extremely difficult (besides variation, inspection results might be conservatively biased); and
- even though it is an affordable inspection approach, it



(a) Dynamic load,  $P$ , as a function of wind speed, adopted from (Sethuraman et al., 2015)



(b)  $c_2$  adjustment factor as a function of contamination  $\eta_c$ , loads  $P_u$ , and viscosity  $\nu$  (SKF-contributors, 2007)

Figure 1. Non-linear time-dependent components of the bearing damage model

is likely that only a limited number of observations are available for training the models (small percentage of the farm can be inspected only a short period of time).

## 3. PROPOSED PHYSICS-INFORMED NEURAL NETWORK

In this paper, we model bearing fatigue (including grease the degradation component) using the concept of hybrid physics-informed neural networks. In such approach, a graph model represent the input-output relationship such that different nodes in the graph can be either physics-based or data-driven nodes. In our implementation, this graph model also repre-

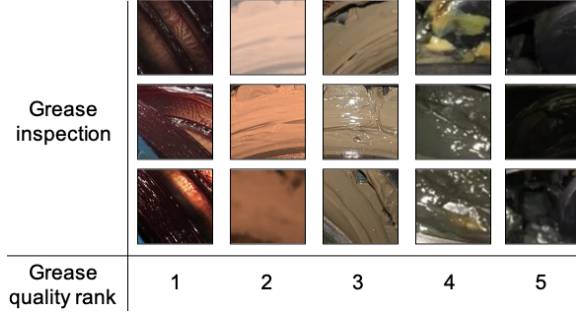


Figure 2. Example of visual inspection ranking system

sents a deep neural network. Given that we are interested in time-dependent bearing fatigue and grease degradation as a function of turbine operation, we use recurrent neural networks (Goodfellow et al., 2016) repeatedly apply transformations to given states in a sequence

$$\mathbf{a}_t = f(\mathbf{x}_t, \mathbf{a}_{t-1}) \quad (2)$$

where  $t \in [0, \dots, T]$  represent the time discretization,  $\mathbf{a}_t = [a_{BRG,t}, a_{GRS,t}]$  are the bearing and grease damage states,  $\mathbf{x}_t$  is the vector of input variables (wind speed and bearing temperature), and  $f(\cdot)$  defines the transition between time steps (function of input variables and previous states).

In this work, we use the Euler integration cell proposed by Nascimento and Viana (2019) and illustrated in Figure 3 to implement numerical integration of Eq. 1. As further reference, Nascimento and Viana (2019); Dourado and Viana (2020); and Yucesan and Viana (2020) demonstrated that the Euler integration cell can be implemented as a hybrid model where data-driven nodes compensate for model-form uncertainty of physics-based nodes.

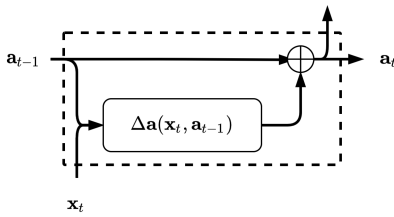


Figure 3. Euler integration recurrent neural network cell

We encapsulate both bearing fatigue damage and the grease damage models in to a recurrent neural network in order to estimate damage accumulation at each cycle, as shown in Figure 4. The recurrent neural network takes wind speed and bearing temperature at each time step (which would come from supervisory control and data acquisition (SCADA) data). Within the cell, there are physics-informed nodes modeling bearing surface fatigue, starting from loads estimation, to grease properties as function of temperature, to life adjustment factors, to finally bearing damage. Given the poorly under-

stood physics of grease degradation, we use a data-driven node to model grease damage increment as a function of current grease damage, wind speed, and bearing temperature. In this paper, we implement this grease damage increment node as a multi-layer perceptron. It can be observed from Figure 4 that we accumulate two damage states, one for grease and one for bearing fatigue, however these quantities can not be observed throughout the operation. Visual grease ranking,  $R_t$ , is the only quantity that helps us to calibrate the data-driven portion of our model.

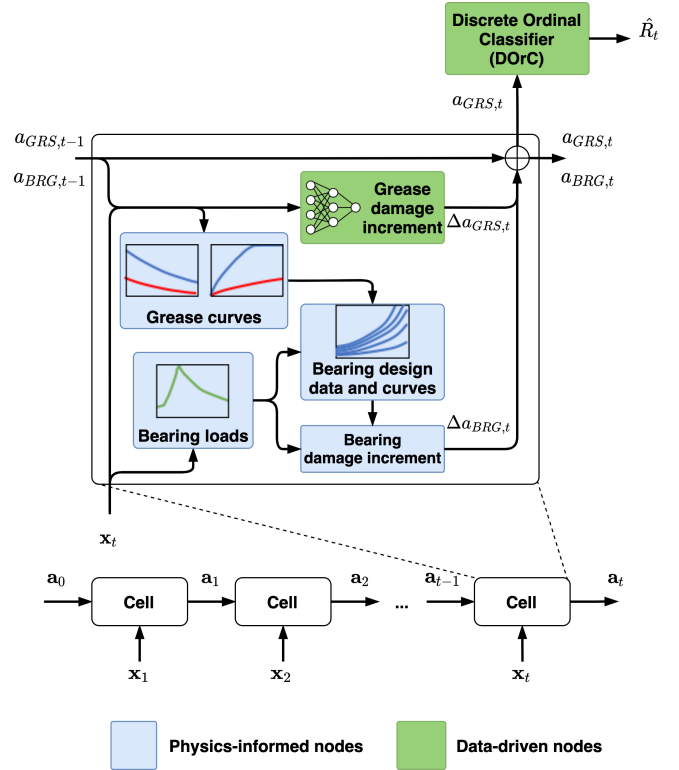
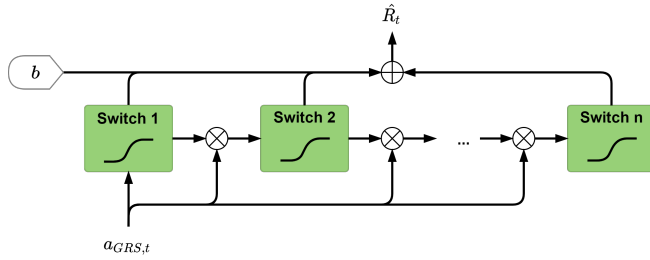


Figure 4. Physics-informed neural network model for bearing fatigue damage and grease damage accumulation

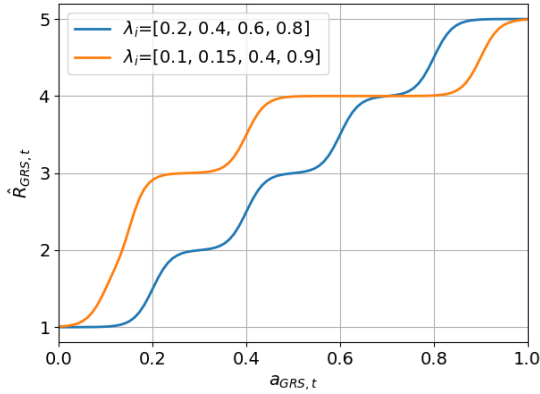
As discussed before in section 2, grease visual inspections returns discrete ratings from 1 to 5, where 1 refers to the pristine state, and 5 refers to the fully degraded state of the grease. Nevertheless, grease damage is a continuous value that starts small when grease is pristine (e.g., 0.0) and increases monotonically throughout the grease useful life (the maximum allowable damage can be normalized to 1.0). Therefore, if the model is limited to the highlighted box of Figure 4, we would not be able to directly compare predicted grease damage  $a_{GRS,t}$  against visual inspection ranking  $R_t$ .

Mapping the predicted grease damage into visual inspection is important. It allows the training of the physics-informed neural network using turbine operational SCADA data as inputs and grease visual inspection as observed output. In order to accomplish that task, **we introduce a novel ordinal classifier**

that we call **discrete ordinal classifier (DOrC)**, shown in Figure 5a. DOrC is a neural network layer that implements a sequence of switches. The first switch takes a scalar as input (predicted grease damage, in our application). The next switch takes the sum of the layer input added to the output of the previous switch as input, and so forth. A parameter  $b$  can shift the final output to the desired lower bound as seen in Figure 5a. In our case, since our ranking is from 1 to 5, we can take  $b = 1$ . As illustrated by Figure 5b, the final input-output relationship resembles a staircase. The major advantage of DOrC, as opposed to simply rounding predictions, is that it can flexibly generate non-linear mappings without the need of specifying the form (e.g. quadratic, cubic, or any other). In fact, during training, the DOrC hyperparameter optimization makes the layer learn the best mapping that represents the observed data.



(a) Graph representation



(b) Examples of DOrC input-output relationship with 4 switches

Figure 5. Discrete ordinal classifier (DOrC)

## 4. CASE STUDY

### 4.1. Turbine Operation Data

In this study, we considered a representative park of 120 wind turbines (detailed in section 2). Site-specific data is obtained from a database provided by NREL (Draxl et al., 2015). This includes environmental data at one hour resolution between 2007 and 2013 for 126,000 different locations throughout the United States. For this case study, we chose a specific area in

Cooke County, TX where an actual wind farm exists. Even though the data does not come directly from the SCADA, we believe using a site where an actual wind farm is located would enhance the similitude of our input data.

Similarly to the procedure adopted by Yucesan and Viana (2020), data is augmented to achieve the 10 minute SCADA resolution; and then, extended up to 30 years to be used for long term bearing fatigue life predictions. Since main bearing temperature is not originally available, we use an analytical model to estimate these values based on ambient and produced power -- see Yucesan and Viana (2020) for further details. Figure 6 illustrates the wind speed and bearing temperature recorded every 10 minutes over 7 years for one turbine.

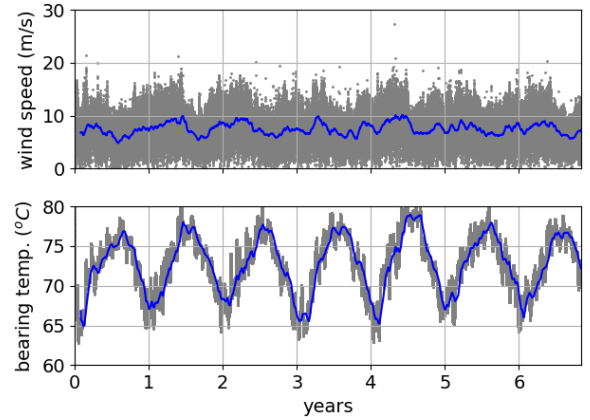


Figure 6. Wind speed and bearing temperature variation of the site over years (grey dots illustrate the actual data, while blue lines are the averaged trends)

### 4.2. Grease and Visual Inspection

The bearing fatigue model needs information about the viscosity and contamination of grease over time. Here, we will scale these two parameters between two assumed known values (one for pristine grease and one for fully degraded grease) using grease damage  $a_{GRS,t}$  as scaling factor:

$$\begin{aligned} \nu_t &= a_{GRS,t}(\nu_{deg} - \nu_{prist}) + \nu_{prist}, \text{ and} \\ \eta_{c,t} &= a_{GRS,t}(\eta_{cdeg} - \eta_{cprist}) + \eta_{cprist}, \end{aligned} \quad (3)$$

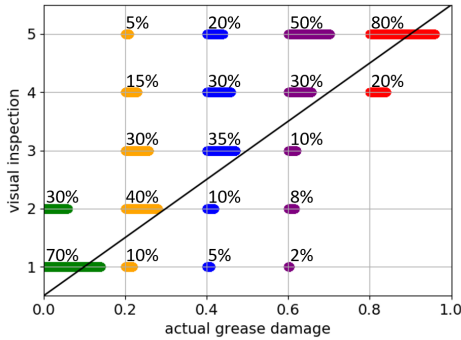
where  $\nu$  and  $\eta_c$  are viscosity and contamination factor of the grease respectively.

In reality, grease damage is as difficult to obtain in real life as accurate values for viscosity and different grease contamination factors. In this study, we will use the models described in Yucesan and Viana (2020) to generate synthetic actual grease damage. It is important to highlight that grease damage information will not be used in the training of our physics-informed neural network. Instead, it will be used to generate the synthetic visual inspection data used in this paper.

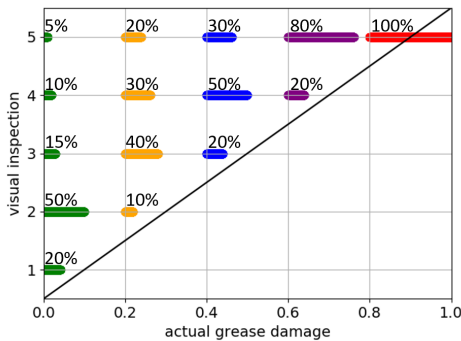
Figure 7 shows two possible scenarios of how variability can

be manifested in visual inspection with probability densities. In both Figures 7a and 7b when actual grease is between 0 and 0.2, the visual inspection results follow the probability distribution shown in green. When actual grease is between 0.2 and 0.4, the visual inspection results follow the probability distribution shown in orange, and so forth. When actual grease is above 0.8, the visual inspection results follow the probability distribution shown in red. This means that for each interval of actual grease, at time of visual inspection, a random ranking between 1 and 5 will be assigned following the corresponding distribution.

We call the scenario shown in Figure 7a ‘‘baseline inspection’’ as the 45° line crosses the 50<sup>th</sup> of each one of the distributions. Even in this scenario, it is important to notice that the distributions are not symmetric. There is probability masses tend to be above the 45° line, and as such, there is a small degree of conservatism in the outcome of the visual inspection. Given that in Figure 7b the probability masses are always above the 45° line (and some are strongly skewed towards the higher ranks), we call it ‘‘conservative inspection’’.



(a) Baseline inspection

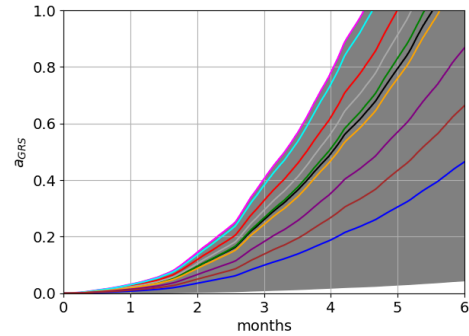


(b) Conservative inspection

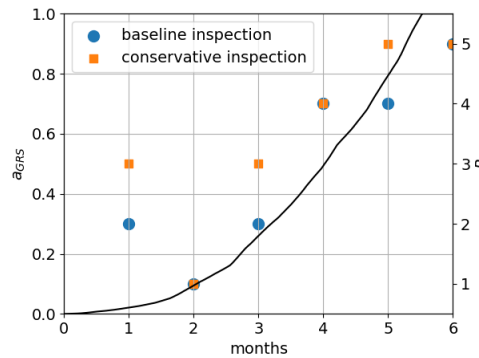
Figure 7. Scenarios of variability in grease visual inspection

We build our grease visual inspection data by assuming that we monitor 10 turbines in the park (out of the 120) for a period of six months. Visual inspection is conducted on a monthly basis. Figure 8a summarizes the grease damage for the 10 turbines used to generate the grease visual inspection data. In

this figure, we also illustrate a shaded region that represent the degradation of entire fleet. Based on this illustration we show that the training turbines mostly consists of aggressive turbines within the farm. In real applications, considering the aggressive portion of a fleet for sampling and training purpose helps the model to accurately predict damages of critical individual machines, and injects conservatism into model (which is preferable in safety assessment applications). We should highlight that the grease damage propagation data shown in Figure 8a is never observed in practice, hence never used in the training of the hybrid model. Here we illustrate the ground true grease damage data (see appendix D for ground true grease degradation data generation) for training turbines which are used as a baseline for ranking sampling presented in Figure 7. Figure 8b exemplifies the outcomes of the visual inspection for one of the turbines used for training. As expected, although grease damage is monotonic, results out of the visual inspection are not necessarily monotonic. This feature is very realistic and we believe it will highlight the robustness of our proposed approach.



(a) Actual (but unknown) grease damage propagation (shaded region indicates grease damage propagation of the entire fleet)



(b) Grease damage and visual inspection for one of the turbines

Figure 8. Grease damage propagation and visual inspection example for the 10 turbines used in the training of the physics-informed neural network

In addition, we considered regreasing operation is performed every 6 months. This will be important later on the paper,

when we make long-term forecast of bearing fatigue based on our physics-informed neural network model. Yucesan and Viana (2019a) studied how to use cumulative damage models to optimize regreasing intervals across a wind park.

### 4.3. Physics-informed Neural Network Design

Given our wind park of 120 turbines, we considered the following information is available for training:

- for every turbine in the park: wind speed and main bearing temperature from SCADA, and
- for 10 turbines in the park: grease visual inspection at every month for six months straight (60 observations in total).

We should note that even though 10 turbines are inspected for training, we are going to predict bearing lives of the entire farm (120 turbines) with the trained model. With that information, we proceed to build the hybrid physics-informed neural network model for bearing fatigue detailed in section 3. In this model, bearing damage accumulation is physics-informed, grease degradation increment,  $\Delta a_{GRS,t}$ , is a multi-layer perceptron, and the mapping between grease damage and visual inspection is done by our proposed discrete ordinal classifier (DOrC).

The configuration of the  $\Delta a_{GRS,t}$  multi-layer perceptron is given in the Table 2. The inputs for this multi-layer perceptron models are scaled between zero and one to avoid that disparities in the order of magnitude of inputs interfere with the fitting of the model.

| Layer | # neurons | activation |
|-------|-----------|------------|
| #1    | 40        | sigmoid    |
| #2    | 20        | elu        |
| #3    | 10        | elu        |
| #4    | 5         | elu        |
| #5    | 1         | sigmoid    |

Table 2. Multi-layer perceptron architecture for grease degradation increment,  $\Delta a_{GRS}$ . Total number of trainable parameters is 1,251 (multi-layer perceptron detailed in the appendix B)

Given that grease visual inspection returns discrete ratings between 1 and 5, our discrete ordinal classifier has four switches. These switches are transition between the ratings and are modeled here as a sigmoid function:

$$\text{switch}_i(x) = \frac{1}{1 + \exp(\alpha(\lambda_i - x))}, \quad (4)$$

where  $i \in [1 \dots 4]$ ,  $\lambda$  is the set of trainable hyperparameters (acting as transition thresholds between ratings), and  $\alpha = -50$  is arbitrarily chosen to make the function steep enough and close to a binary transition (while smooth enough to avoid discontinuities during training of the deep neural network). This

way, by adjusting each threshold, we can train our classifier to map the given continuous damage index to the discrete rank scale. Although we let thresholds to be learned by the model, we imposed the following bounds:

$$\begin{aligned} \lambda_1 &= [0.0, 0.3], & \lambda_2 &= [0.2, 0.5], \\ \lambda_3 &= [0.4, 0.8], & \text{and } \lambda_4 &= [0.8, 2.0]. \end{aligned} \quad (5)$$

Here, we used the mean squared error as the loss function while optimizing the trainable parameters of the stacked recurrent neural network (physics-informed neural network and DOrC layer):

$$\text{Loss} = \frac{1}{N_O} \sum_j \sum_i \left( R_{ij}^{GRS} - \hat{R}_{ij}^{GRS} \right)^2, \quad (6)$$

where  $N_O$  is the total number of observations,  $R_{ij}^{GRS}$  is the  $i^{th}$  observation of grease damage rank for  $j^{th}$  turbine, and  $\hat{R}_{ij}^{GRS}$  is the predicted grease damage rank for the  $i^{th}$  grease visual sample of the  $j^{th}$  turbine.

Optimizing the 1,251 (multi-layer perceptron) + 4 (DOrC thresholds) trainable parameters can be a challenging task. An initial point far away from actual relationship might cause divergence or very long time of training process. Therefore, initializing the weights and biases of this neural network model can greatly improve the training process. We follow the same strategy presented by Yucesan and Viana (2020) and summarized in the appendix A. After weights are initialized, we used RMSprop<sup>1</sup> set with learning rate 0.0005 and 2500 epochs. Overall algorithm flowchart for data collection, training, and predicting is as shown in Figure 9.

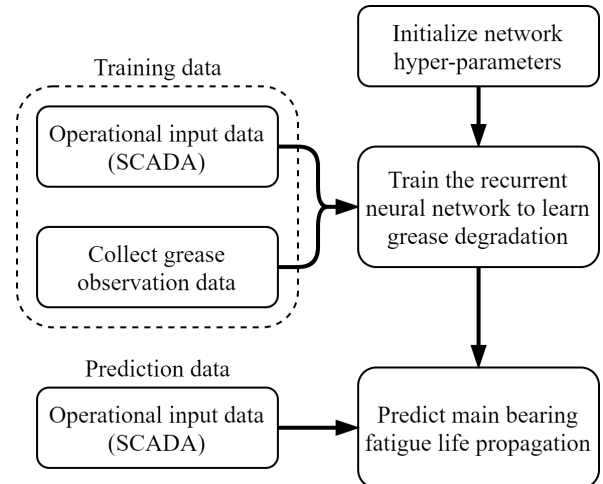


Figure 9. Algorithm block diagram for the entire training and prediction approach

<sup>1</sup>[www.tensorflow.org/api\\_docs/python/tf/keras/optimizers/RMSprop](http://www.tensorflow.org/api_docs/python/tf/keras/optimizers/RMSprop)

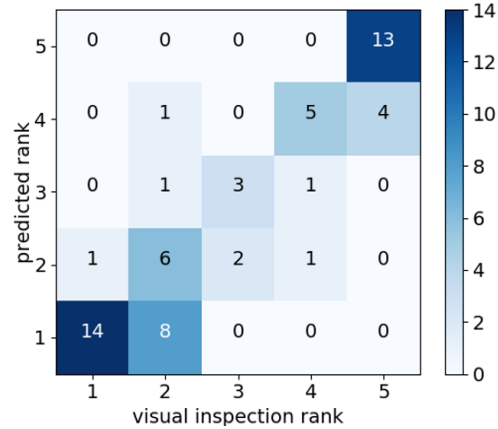
#### 4.4. Replication of results

Our implementation is done in TensorFlow (version 2.0.0-beta1) using the Python application programming interface. In order to replicate our results, the interested reader can download codes and data. First, install the PINN package (base package for physics-informed neural networks used in this work) available at Viana et al. (2019). Then, clone the ‘‘pinn\_wind\_bearing’’ repository found in Yucesan and Viana (2019b) and go to folder ‘‘phm\_2020’’. This repository includes three python scripts where the first one samples visual grease inspections based on ground true grease data, the second one trains the recurrent neural network using a pretrained multi-layer perceptron model with fixed initial weights, and the last script predicts the fatigue damage accumulation of the wind turbine main bearing for 20 years. The reason we limited the time frame to 20 years and not 30 years as we used in the paper, that the size limitation of the database we use to share our data<sup>2</sup>. The data used in this work is publicly available in Yucesan (2020). Download the data and extract folders inside ‘‘wind\_bearing\_dataset\_2020’’ to the directory where the ‘‘pinn\_wind\_bearing/phm\_2020’’ repository is cloned. All simulations were conducted using a laptop configured with an Intel Core i7-8650U CPU at 1.90GHz, 32GB of RAM, and NVIDIA Quadro P500 graphical processing unit running Windows 10.

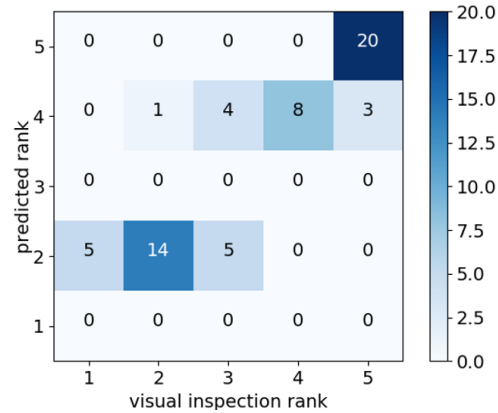
#### 5. RESULTS AND DISCUSSION

We start by analyzing the results out of the training of our stacked physics-informed recurrent neural network. We use data out of 10 turbines, where input data is the wind speed and main bearing temperature in 10-minute intervals and output is the grease visual inspection. The data is used to simultaneously optimize for the network hyperparameters, 1,251 from the multi-layer perceptron and 4 from the DOrC layer. Figure 10 shows the confusion matrices out of the predictions coming from networks trained with both the baseline and conservative inspection data. Given that the training of the neural network uses the mean square error as loss function, in both cases, the networks will result in unbiased predictors. The caveat is that these are unbiased predictors for the grease visual inspection. Unfortunately, as we discussed in section, grease visual inspection is prone to large uncertainty (with both bias and variance). Therefore, when we compare the predicted and actual grease damage (instead of the raking from grease visual inspection), we should be able to see the manifestation of such uncertainty in predictions.

Figure 11 shows the comparison between predicted rank (out of the trained stacked recurrent neural networks) and actual grease damage grouped in bins similarly to the ones of Figure 7. It confirms that, while predicted ranks are unbiased for models trained with data coming from both baseline and con-



(a) Baseline inspection



(b) Conservative inspection

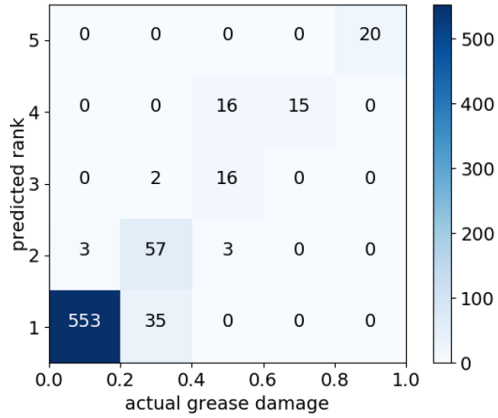
Figure 10. Confusion matrices of the training set

servative visual inspection scenarios, there is significant bias introduced by the conservative rankings. As expected, predictions coming out of the model trained with baseline inspection (Figure 11a) tend to be around the diagonal of the matrix, even though there are deviations. Likewise, predictions coming out of the model trained with conservative inspection (Figure 11b) are mostly above the diagonal of the matrix. These numbers are realizations of the probability distributions shown in Figure 7.

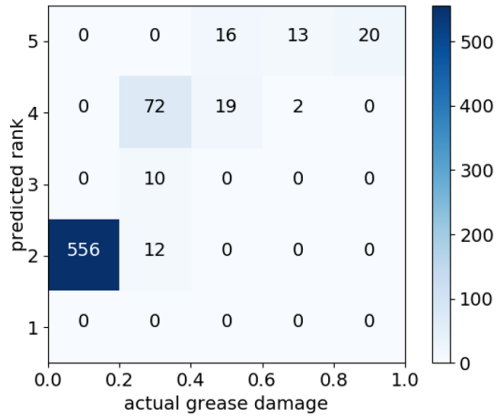
One benefit of building a stacked recurrent neural network with the physics-informed neural network and DOrC layer is that, once the model is trained, we can use the physics-informed neural network to estimate grease damage. This is possible, even though actual grease damage was never observed and the model was trained with only grease visual inspection. Figure 12 presents the prediction results at the training set after models are trained. These time history predictions of grease damage should be compared against Figure 8a. As expected, while the training using baseline inspection yields to relatively accurate predictions (Figure 12a), training using conservative inspection scenario estimates grease damage accumulation at a rate higher than the actual one (Fig-

<sup>2</sup><https://dataverse.harvard.edu/>





(a) Baseline inspection

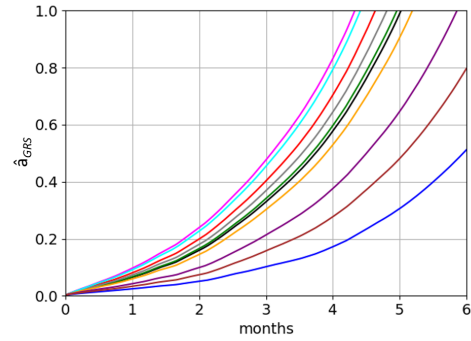


(b) Conservative inspection

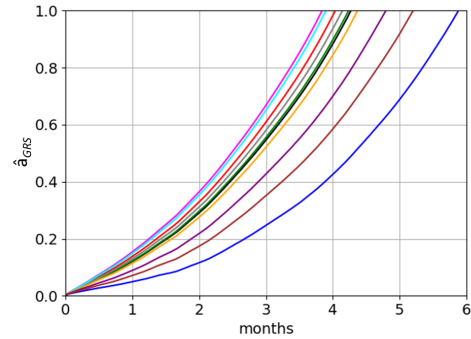
Figure 11. Predicted grease quality rank versus actual grease damage across entire wind park (120 turbines)

ure 12b). Figure 13 illustrates how these predicted grease damage compares against the actual (but unknown) grease damage across the entire wind park. The model trained with the baseline inspection results is significantly less conservative (0.3% RMSE) than the one trained with the conservative inspection results (5% RMSE). We understand most practitioners would recognize that visual inspection can be biased. Nevertheless, we believe that, in real life, the degree of bias (conservatism) would not be known. In future research, we want to investigate ways to quantify and reduce prediction bias when our physics-informed neural networks are trained with biased observations.

Finally, we used the physics-informed neural network model to estimate both grease damage accumulation and main bearing fatigue damage. Figure 14 illustrates the results for one turbine of the park that was not in the training set. Figure 14a shows the grease damage prediction results of our neural network predictions and actual grease damage over time. While model trained with conservative inspections overshoots the actual propagation in every period, model trained with baseline inspections tends to follow the trend in relatively good



(a) Baseline inspection



(b) Conservative inspection

Figure 12. Predictions of grease damage propagation for the 10 turbines used in the training of the physics-informed neural network models

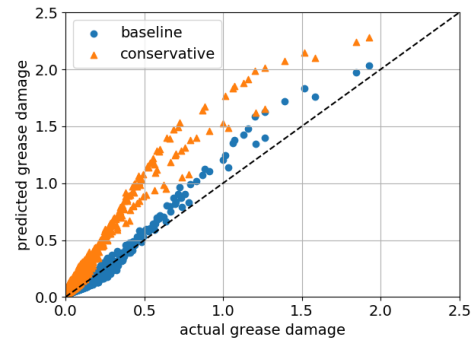
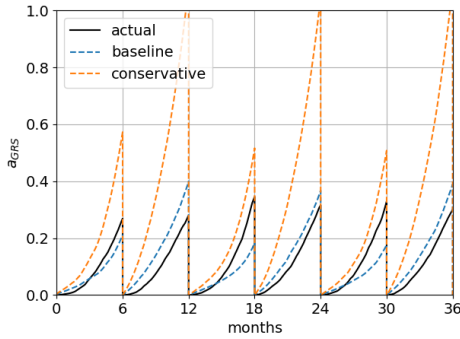


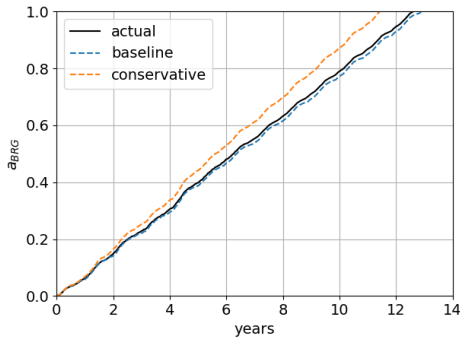
Figure 13. Grease damage predictions vs. actual grease damage for the wind park (120 turbines)

agreement (albeit there is still small conservatism). Figure 14b presents the bearing fatigue damage prediction results of our neural network predictions and actual bearing fatigue damage over time. Even though there are different degrees of conservatism in grease damage estimation, the bearing fatigue damage estimation is in relatively good agreement. Bearing fatigue damage seems to be only marginally overestimated (conservatism), even when models are trained with conservative grease inspection. The reason for such behavior is the

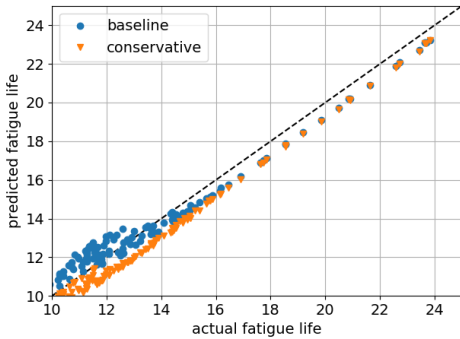
regreasing policy. In this study, we assume that bearings are fully regreased every six months. Therefore, unless discrepancies in grease damage are substantially large for most of those six months, there would not be significant discrepancies in bearing fatigue damage. Figure 14c summarizes the bearing fatigue damage estimation results showing the time-to-failure (i.e., time needed for bearing fatigue damage to reach 1.0) for the wind park. Interestingly, for the model trained with the baseline grease inspections, there is considerable scatter in the earlier failures of the farm, specially between 10 and 13 years. After that, both models tend to be conservative (and are equally conservative after for failures happening 16 years).



(a) Grease damage propagation of a sample turbine



(b) Bearing fatigue damage propagation of a sample turbine



(c) Predicted against actual bearing fatigue life across wind park

Figure 14. Grease and bearing fatigue damage predictions

## 6. SUMMARY AND FUTURE WORK

In this study, we modeled the fatigue damage accumulation of the main bearing component of the wind turbine with a hybrid physics-informed neural networks approach. While we modeled the fatigue damage propagation with physics-based relations, the grease damage increment is represented by neural networks. The main challenge addressed here was the estimation of missing physics using only turbine operation data as input and grease visual inspection as output. In order to achieve that, we constructed a custom classifier to map continuous grease damage scale into discrete ranks. This allowed the model to be robust to uncertainties due to visual grease inspection routine.

In the case study used to illustrate the capabilities of our framework, we considered that in a wind park of 120 turbines, 10 were inspected every month for 6 months. We also considered inspections with different levels of uncertainty. Results from the case study showed that our physics-informed neural networks model can simultaneously learn the grease damage accumulation and the classification. We also learned that what we called baseline inspection (in which ranking distribution is skewed but 50 percentile follows a linear relationship with actual damage) lead to model that successfully estimated grease damage accumulation and eventually accurately predicted bearing fatigue damage accumulation. However, when visual inspection is conservative (highly skewed ranking distribution), the resulting model predictions are also conservative for both grease and bearing damage. This behavior was expected and illustrates why wind park operators tend to be extra cautious when performing visual inspection.

Finally, in the light of the results obtained thus far, we also would like to extend this study by including following items as potential future research:

- accounting for the uncertainty within the inputs (i.e. uncertainty in the loads model and material capabilities) and effect of number of samples in the model's performance,
- exploring the abilities of the model for operational benefits, such as fleet recommissioning (life extension), bearing replacement, inspection scheduling and other financial savings.

## ACKNOWLEDGMENT

This work was supported by the University of Central Florida (UCF). Nevertheless, any view, opinion, findings and conclusions or recommendations expressed in this material are those of the authors alone. Therefore, UCF does not accept any liability in regard thereto.

## NOMENCLATURE

|                  |   |
|------------------|---|
| $a_{BRG}$        | cumulative bearing fatigue damage                 |
| $a_{GRS}$        | cumulative grease damage                          |
| $\Delta a_{BRG}$ | incremental bearing fatigue damage                |
| $\Delta a_{GRS}$ | incremental grease damage                         |
| $P$              | equivalent dynamic bearing load                   |
| $P_u$            | fatigue load limit                                |
| $C$              | design load rating                                |
| $c_1$            | reliability level factor                          |
| $c_2$            | life adjustment factor                            |
| $V_W$            | wind speed  |
| $T_{BRG}$        | bearing temperature                               |
| $\nu$            | viscosity   |
| $\eta_c$         | contamination factor                              |
| $R$              | grease visual inspection rank                     |
| $b$              | discrete ordinal classifier shifting constant     |
| $\alpha$         | discrete ordinal classifier steepness coefficient |
| $\lambda$        | discrete ordinal classifier switch threshold      |

## REFERENCES

- Bergs, C., Heizmann, M., & Held, H. (2018, July). Hybrid modeling approaches with a view to model output prediction for industrial applications. In *2018 IEEE 16th international conference on industrial informatics (INDIN)*. IEEE. Retrieved from <https://doi.org/10.1109/indin.2018.8471964> doi: 10.1109/indin.2018.8471964
- Cambron, P., Tahan, A., Masson, C., & Pelletier, F. (2017). Bearing temperature monitoring of a wind turbine using physics-based model. *Journal of Quality in Maintenance Engineering*, 23(4), 479–488.
- Dourado, A., & Viana, F. A. C. (2020). Physics-informed neural networks for missing physics estimation in cumulative damage models: a case study in corrosion fatigue. *ASME Journal of Computing and Information Science in Engineering*, 20(6), 061007 (10 pages). Retrieved from <https://doi.org/10.1115/1.4047173> doi: 10.1115/1.4047173
- Draxl, C., Clifton, A., Hodge, B.-M., & McCaa, J. (2015). The wind integration national dataset (wind) toolkit. *Applied Energy*, 151(1), 355–366.
- GE-contributors. (2009). *GE Energy 1.5 MW Wind Turbine Brochure*. Online (retrieved 23 May 2018). Retrieved from <https://geosci.uchicago.edu/~moyer/GEOS24705/Readings/GEA14954C15-MW-Broch.pdf>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. Retrieved from <http://www.deeplearningbook.org>
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. doi: 10.1162/neco.1997.9.8.1735
- Hornemann, M., & Crowther, A. (2013). *Establishing failure modes for bearings in wind turbines*. Online (retrieved 23 May 2018). Retrieved from <https://www.windpowerengineering.com/establishing-failure-modes-for-bearings-in-wind-turbines/>
- Klueber-contributors. (2011). *The element that rolls the bearing*. Online (retrieved 2 December 2018). Retrieved from [https://www.klueber.com/ecomaXL/files/Competence\\_brochure\\_rolling\\_bearing%5B1%5D.pdf](https://www.klueber.com/ecomaXL/files/Competence_brochure_rolling_bearing%5B1%5D.pdf)
- Lugt, P. M. (2009). A review on grease lubrication in rolling bearings. *Tribology Transactions*, 52(4), 470–480.
- Nascimento, R. G., & Viana, F. A. C. (2019, September). Fleet prognosis with physics-informed recurrent neural networks. In *12th International Workshop on Structural Health Monitoring* (pp. 1740–1747). Stanford, USA. doi: 10.12783/shm2019/32301
- Raissi, M. (2018). Deep hidden physics models: deep learning of nonlinear partial differential equations. *Journal of Machine Learning Research*, 19(25), 1–24. Retrieved from <http://jmlr.org/papers/v19/18-046.html>
- Sethuraman, L., Guo, Y., & Sheng, S. (2015). Main bearing dynamics in three-point suspension drivetrains for wind turbines. In *American wind energy association conference & exhibition*. Orlando, USA: AWEA.
- SKF-contributors. (2007). *SKF spherical roller bearings catalogue*. Online (retrieved 5 June 2018). Retrieved from [http://www.skf.com/binary/30-148465/6100\\_EN.pdf](http://www.skf.com/binary/30-148465/6100_EN.pdf)
- Viana, F. A. C., Nascimento, R. G., Yucesan, Y., & Dourado, A. (2019, Aug). *Physics-informed neural networks package*. <https://github.com/PML-UCF/pinn>. Zenodo. Retrieved from <https://github.com/PML-UCF/pinn> doi: 10.5281/zenodo.3356877
- Walker, C., & Coble, J. (2018). Wind turbine bearing fault detection using adaptive resampling and order tracking. *International Journal of Prognostics and Health Management*, 9(2).
- Watanabe, F., & Uchida, T. (2015). Micro-siting of wind turbine in complex terrain: simplified fatigue life prediction of main bearing in direct drive wind turbines. *Wind Engineering*, 39(4), 349–368.
- Yucesan, Y. A. (2020). *Replication data for: A hybrid modeling for wind turbine main bearing fatigue with uncertainty in grease observations*. Harvard Dataverse. Retrieved from <https://doi.org/10.7910/DVN/J3IUVR> doi: <https://doi.org/10.7910/DVN/J3IUVR>
- Yucesan, Y. A., & Viana, F. A. C. (2019a, Jan). Onshore wind turbine main bearing reliability and its implications in fleet management. In *AIAA Scitech 2019 Forum* (p. AIAA-2019-1225). Orlando, USA: AIAA. Retrieved from <http://dx.doi.org/10.2514/>

6.2019-1225 doi: 10.2514/6.2019-1225

Yucesan, Y. A., & Viana, F. A. C. (2019b, Aug). *Python scripts for wind turbine main bearing fatigue life estimation with physics-informed neural networks*. Zenodo. Retrieved from [https://github.com/PML-UCF/pinn\\_wind\\_bearing](https://github.com/PML-UCF/pinn_wind_bearing) doi: 10.5281/zenodo.3355725

Yucesan, Y. A., & Viana, F. A. C. (2020). A physics-informed neural network for wind turbine main bearing fatigue. *International Journal of Prognostics and Health Management*, 11(1), 27-44. Retrieved from <http://www.phmsociety.org/node/2736>

Zhu, J., Yoon, J. M., He, D., Qu, Y., & Bechhoefer, E. (2013). Lubrication oil condition monitoring and remaining useful life prediction with particle filtering. *International Journal of Prognostics and Health Management*, 4, 124-138.

## BIOGRAPHIES

**Yigit A. Yucesan** received his B.S. degree in Aerospace Engineering from the Middle East Technical University in 2015. After obtaining his B.S. degree, he worked as a Structural Analysis and Methods Engineer at Turkish Aerospace Industries (TAI) for 2.5 years, and in the mean time he completed his M.S. degree in Mechanical Engineering from TOBB University of Economics and Technology in 2018. He is currently a Graduate Research Assistant in the University of Central Florida and a Ph. D. candidate in Mechanical Engineering focusing on physics-informed neural networks for wind turbine main bearing fatigue modeling.

**Felipe A. C. Viana** is currently an assistant professor at the University of Central Florida (UCF). The vast majority of Dr. Viana's work has been applied to new designs and improvement of fielded products with a focus on aircraft propulsion, power generation, and oil and gas systems. Before joining UCF, Dr. Viana was a Sr. Scientist at GE Renewable Energy, where he led the development of state-of-the-art computational methods for improving wind energy asset performance and reliability. Prior to moving to that role at GE, he spent five years at GE Global Research, where he led and conducted research on design and optimization under uncertainty, probabilistic analysis of engineering systems, and services engineering.

## Appendix A: Physics-guided neural network weight initialization

We understand that training deep recurrent neural networks is challenging due to the high number of hyperparameters and nonlinear output behavior. Therefore, we advocate for initializing the data-driven nodes in the model whenever possible. In this work, similarly to (Yucesan & Viana, 2020), we suggest initializing the grease damage increment model by making it conform to a simple linear plane representation with the

following input-output relationship:

$$\Delta a_{GRS} = \alpha_0 + \alpha_1 \times T_{BRG} + \alpha_2 \times V_W + \alpha_3 \times a_{GRS}, \quad (7)$$

where  $\Delta a_{GRS}$  is the grease damage increment,  $T_{BRG}$  is the main bearing temperature,  $V_W$  is the wind speed, and  $a_{GRS}$  is predicted cumulative grease damage.

The coefficients,  $\alpha_i$ , are initialized using engineering judgment. For example, we can safely assume that  $\Delta a_{GRS}$  increases with increasing bearing temperature; therefore, the  $\alpha_1$  has to be positive. Similarly, engineering judgment can be used to limit  $\Delta a_{GRS}$ , which is expected to be on the order of magnitude of the observed  $a_{GRS}$  divided by the number of time intervals (i.e., cycles). For illustration purpose, one of the randomly generated plane is plotted against the actual input output relationship in Figure 15. In this illustration, wind speed and bearing temperature are the two inputs of multi-layer perceptron (with  $a_{GRS}$  fixed at 0.5) and the grease damage increment  $\Delta a_{GRS}$  is the output of the multi-layer perceptron. The orange surface in the plot represents the actual (but unknown) input output behavior and the blue plane is the approximation to this behavior given by the multi-layer perceptron.

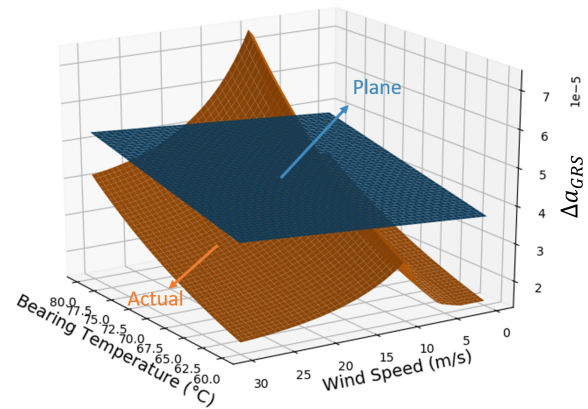


Figure 15. Plane approximation to actual data (in this illustration,  $a_{GRS} = 0.5$ )

## Appendix B: Multi-layer perceptron

In this study we used a multi-layer perceptron to model incremental grease damage output. A multi-layer perceptron consists of multiple layers with different numbers of neurons. Each neuron has a weight vector the same length as the inputs going into that neuron, and optionally a bias term. After the inputs are multiplied by weights and a bias is added, the result of this operation is fed into the activation function of the neuron as input, and yields to output of the neuron. For example, for a single neuron with sigmoid activation function, the formulation becomes:

$$y = \text{sgm}(\mathbf{w}^T \mathbf{x} + b) \quad (8)$$

where  $w$  and  $b$  are trainable hyperparameters. Table 3 presents a three layered multi-layer perceptron with sigmoid and exponential linear unit (elu) activation functions. For this architecture, the diagram of the multi-layer perceptron that takes two inputs ( $x_1, x_2$ ) and gives one output ( $y$ ), is provided in Figure 16, and the activation functions can be found in Eq. 9.

| Layer | # neurons | activation |
|-------|-----------|------------|
| #1    | 3         | sigmoid    |
| #2    | 2         | elu        |
| #3    | 1         | sigmoid    |

Table 3. Example architecture for a simple multi-layer perceptron architecture

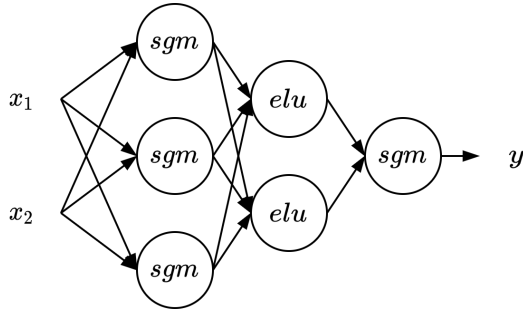


Figure 16. Flow diagram of a multi-layer perceptron

$$\text{sgm}(z) = \frac{1}{1 + e^{-z}},$$

$$\text{elu}(z) = \begin{cases} z & \text{when } z > 0, \text{ and} \\ e^z - 1 & \text{otherwise.} \end{cases} \quad (9)$$

### Appendix C: Study of pure data-driven models

In this appendix, we compare our hybrid physics-informed neural network approach to a conventional pure data-driven model: long short-term memory (LSTM) recurrent neural network cell (Hochreiter & Schmidhuber, 1997).

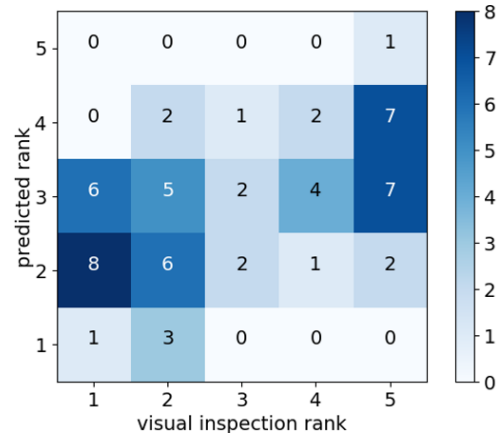
We chose two different complexity levels for LSTM cells: one is a single layer architecture, that we call "Shallow LSTM", and the other one consists of multiple layers, we call "Deep LSTM". Table 4 summarizes the architectural details for these models.

We trained both models with the same optimization settings and training data used to train our hybrid model discussed in the section 4.3. Figure 17 present the confusion matrices after training of both models. It is fair to say that pure data-driven models exhibit decent performance in predicting the noisy visual inspections, and the complexity level only makes marginal difference in the prediction performance. However, we suspect these models tend to fit the data by disregarding

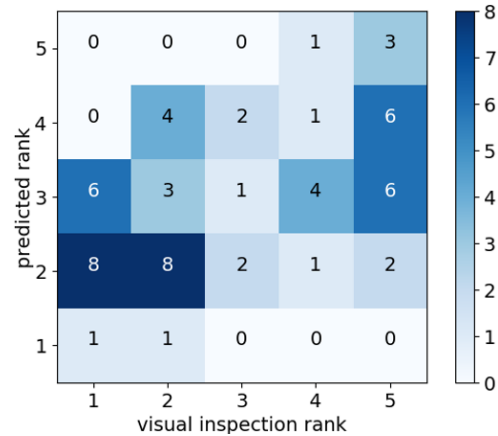
Table 4. Long short-term memory (LSTM) network designs.

| Design     | Shallow LSTM | Deep LSTM |
|------------|--------------|-----------|
| Layers     | 1            | 3         |
| Neurons    | 8            | 8         |
| Parameters | 360          | 1,448     |

the ordinal nature of the problem. In fact, Figure 18 proves this point, as it illustrates the time history prediction of LSTM models against our physics-informed neural network approach for a single turbine within the training set. Not only LSTM models perform poorly to approximate time history prediction of grease ranks, they don't physically make sense (as predictions go up and down). On the other hand, the hybrid approach we proposed with a novel classifier DORC, adopts the damage accumulation phenomenon thanks to physics-informed nodes, and performs very well to predict grease rank evolution over time.



(a) Shallow LSTM



(b) Deep LSTM

Figure 17. Confusion matrices obtained on the training set (predicted ranks versus ranks given by visual inspection)

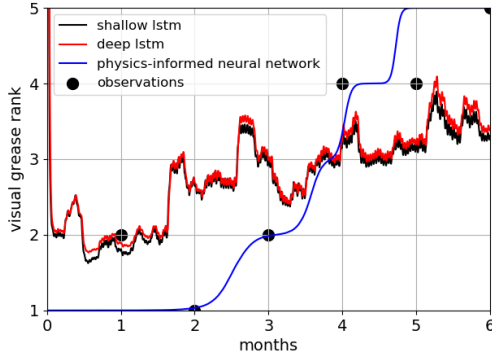


Figure 18. Ranking prediction from recurrent neural networks versus observed rank for one of the turbines in the training set.

#### Appendix D: Baseline grease degradation data

Grease degradation is a complex phenomenon to model. In this paper, we adopted a simplified model found in (Klueber-contributors, 2011) to form our baseline ground true data for grease degradation. The model relates grease life with bearing temperature and a number of adjustment factors:

$$L_{nm}^{GRS} = L_{nm}^{GRS*} K_N K_B F_1 F_2 F_3 F_4 F_5 F_6 \quad (10)$$

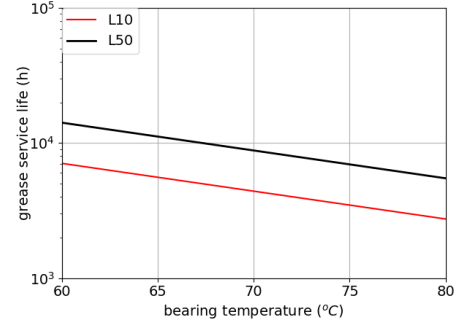
Figure 19a illustrates how grease service life varies with temperature. Most adjustment factors are given in Table 5.  $F_3$  is a factor that accounts for dynamic load variation and it is shown in Figure 19b. As stated by Lugt (2009), the bearing life is commonly expressed in terms of  $L_{10}$  life (as a safety factor to account for the variation in grease properties).

| Parameter | Value | Account for                       |
|-----------|-------|-----------------------------------|
| $K_N$     | 7.69  | Bearing design                    |
| $K_B$     | 0.15  | Spherical bearing design          |
| $F_1$     | 0.8   | Dust and humidity                 |
| $F_2$     | 0.9   | Shock, vibration, and oscillation |
| $F_4$     | 1.0   | Air flow                          |
| $F_5$     | 1.0   | Rotating outer ring               |
| $F_6$     | 1.0   | Vertical shaft arrangement        |

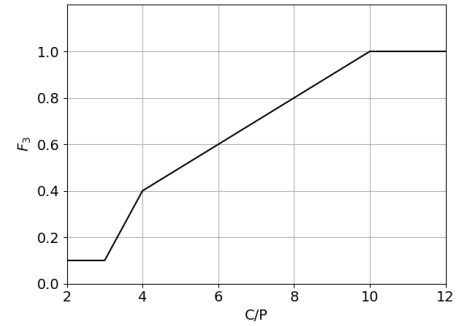
Table 5. Grease modification factors adopted from Klueber-contributors (2011)

#### Appendix E: Input data preprocessing

In wind turbines, supervisory control and data acquisition (SCADA) systems are usually available on board. SCADA systems record data from sensors and control system every 10 minutes. In this study, we assume wind speed and main bearing temperature are provided through SCADA system for every turbine of the fleet. However, the data we can extract from NREL database is wind speed and ambient temperature at 80 meters altitude recorded every hour. In order to represent SCADA data, we bootstrapped data obtained from NREL



(a) Nominal grease service life versus bearing temperature



(b) Grease life adjustment factor depending on the dynamic load

Figure 19. Grease life and  $F_3$  adjustment factor adopted from Klueber-contributors (2011).

database. Each day is represented by eight bins of three hours segments and each bin aggregates a week worth of data. In other words, each bin has 21 data points coming from the same 3 hours of the day across a week. We then sample at random (with replacement) from this pool to fill in the extra 5 points per hour needed within each bin. This process is repeated with a sliding weekly window throughout the year so that seasonality is preserved. While the NREL database covers 7 years, some of our simulations needed data for up to 30 years. To overcome this limitation and also to provide a mechanism for forecasting damage accumulation. Again, we bootstrapped from the previously augmented data binning it at every ten minutes by time of the day and day of the year across the seven years. We calculated the mean and standard deviation of each bin and assuming normal distribution, we sampled data points for the same time stamp of the forecasted year.

As we mentioned before, the NREL database provides ambient temperature, however our model requires main bearing temperature. In order to preprocess the temperature data, we used the model proposed by Cambron et al. (2017). In essence, the main bearing temperature is described by a recursive model as a function of previous bearing temperature, nacelle temperature, angular velocity, and generated power.