# D5.2 Final Report - Enabling Condition Based Maintenance (CBM) for Offshore Wind

Prepared for:

**New York State Energy Research and Development Authority**

Albany, NY

and

**The National Offshore Wind Research & Development Consortium**

Albany, NY

Melanie Schultz
Program Manager

Prepared by:

**GE Research**

Niskayuna, NY

Project Manager 1: Guiju Song, PhD
Offshore Wind Platform Leader at GE Research

Project Manager 2: Susan Kocsmiersky
Contract Administrator at GE Research

# Notice

This report was prepared by Dr. Guiju Song in the course of performing work contracted for and sponsored by the New York State Energy Research and Development Authority (hereafter "NYSERDA"). The opinions expressed in this report do not necessarily reflect those of NYSERDA or the State of New York, and reference to any specific product, service, process, or method does not constitute an implied or expressed recommendation or endorsement of it. Further, NYSERDA, the State of New York, and the contractor make no warranties or representations, expressed or implied, as to the fitness for particular purpose or merchantability of any product, apparatus, or service, or the usefulness, completeness, or accuracy of any processes, methods, or other information contained, described, disclosed, or referred to in this report. NYSERDA, the State of New York, and the contractor make no representation that the use of any product, apparatus, process, method, or other information will not infringe privately owned rights and will assume no liability for any loss, injury, or damage resulting from, or occurring in connection with, the use of information contained, described, disclosed, or referred to in this report.

NYSERDA makes every effort to provide accurate information about copyright owners and related matters in the reports we publish. Contractors are responsible for determining and satisfying copyright or other use restrictions regarding the content of reports that they write, in compliance with NYSERDA's policies and federal law. If you are the copyright owner and believe a NYSERDA report has not properly attributed your work to you or has used it without permission, please email print@nyserda.ny.gov

Information contained in this document, such as web page addresses, are current at the time of publication.

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Wind Energy Technologies Office Award Number DE-EE0008390.

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily

constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

# Table of Contents

**List of Figures** ..... **6**

**List of Tables** ..... **7**

**Acronyms and Abbreviations** ..... **8**

**Executive Summary** ..... **9**

**1 Task 1: Generating data using physical model/simulator** ..... **11**

1.1 Completion simulator preparation ..... 11

1.2 Simulation data generation ..... 11

1.3 Observability study completion ..... 12

1.4 Discussions of the observations and findings from Task 1 ..... 13

**2 Task 2: Building health index estimation (HIE) model** ..... **14**

2.1 Completion of feature selection for HIE model ..... 14

2.1.1 Model feature selection ..... 14

2.1.2 Model feature selection observations, findings, and recommendations ..... 16

2.2 Health index estimation (HIE) modeling ..... 19

2.2.1 HIE modeling strategies and techniques ..... 19

2.2.2 Model performance metrics and evaluation methods ..... 22

2.2.3 HIE modeling results ..... 22

2.3 Discussions of the observations and findings from Task 2 ..... 26

**3 Task 3: Calibrating HIE model using the SCADA data** ..... **27**

3.1 Completion of filed data preparation and calibration strategy ..... 27

3.1.1 Motivation for calibration ..... 27

3.1.2 Model calibration strategy ..... 28

3.2 Completion of HIE model calibration and validation ..... 28

3.2.1 Apply model calibration to field data ..... 28

3.2.2 Scenario 1 – "Calibration via re-simulation" ..... 29

3.2.3 Scenario 2 – Calibration via model updating ..... 36

3.3 Discussions of the observations and findings from Task 3 ..... 41

**4 Task 4: Building RUL prediction mode** ..... **43**

4.1 A suitable roughness growth model identified ..... 43

4.1.1 State of the art prognostics methods ..... 43

4.1.2 Challenges in in building blade roughness growth model for offshore wind ..... 46

4.1.3 Blade LEE growth model and validation approaches ..... 48

4.2 Completion of the initial prediction model and data preparation; Completion of RUL model building and validation ..... 49
4.2.1 Data generation for RUL model ..... 49
4.2.2 RUL model development and validation ..... 55
4.3 Discussions of the observations and findings from Task 4 ..... 60
**5 Project summary ..... 61**
**6 Bibliography ..... 62**

# List of Figures

Figure 1. Parameters and levels used in initial DOE ..... 12
Figure 2. Simulation file components ..... 12
Figure 3: Scatter plot between the true and the predicted roughness levels. ..... 25
Figure 4: Histograms of the predicted roughness levels. ..... 25
Figure 5. Calibration steps. ..... 28
Figure 6. Environmental condition comparison for DOE and 12MW field data. ..... 30
Figure 7 Calibration via re-simulation. ..... 31
Figure 8. 3D view of the environmental condition for DOE and DOE3a. ..... 32
Figure 9. Environmental condition comparison for DOE, DOE3a and 12MW field data. ..... 32
Figure 10. Scatter plot between the true and the predicted roughness levels on simulated data. ..... 33
Figure 11. Histograms of the predicted roughness levels (for 12MW simulation data). ..... 34
Figure 12. Initial model performance on field data using DOE (for 12MW). ..... 35
Figure 13. Model adaption performance using DOE3a (for 12MW) on field data. ..... 36
Figure 14. Calibration via model updating. ..... 37
Figure 15. Model adaption using DOE3a performance on 14MW field data. ..... 38
Figure 16. Scatter plot between the true and the predicted roughness levels on simulation data (for 14MW). ..... 39
Figure 17. Histograms of the predicted roughness levels (for 14MW) on simulated data. ..... 39
Figure 18. Model adaption performance using DOE3a and 50 samples of 14MW field data. ..... 40
Figure 19. Model adaption performance using DOE3a and 50 samples of 14MW field data. ..... 41
Figure 20. Various sources of uncertainties affecting accuracy of prognostics (adapted from [5]). ..... 46
Figure 21. Illustrative roughness growth trends using simulation ..... 48
Figure 22. Relationship between roughness and AEP loss ..... 50
Figure 23. Lab test data ..... 51
Figure 24. Baseline RUL model ..... 52
Figure 25. Precipitation at Rotterdam ..... 53
Figure 26. Data generation summary ..... 54
Figure 27. An RUL model building example ..... 56
Figure 28. Model validation for LEP ..... 58
Figure 29. Model validation for no-LEP ..... 59

# List of Tables

Table 1. Haliade X descriptive and performance parameters .....11
Table 2. Forward stepwise selection algorithm. .....15
Table 3: RF algorithm. .....16
Table 4: Variable selection comparison. .....17
Table 5: Variable selection comparison for selected wind speed range. .....18
Table 6: Input variables of our HIE model. .....23
Table 7: HIE model performance comparison. .....26
Table 8. HIE model performance on simulated data (train and test). .....34
Table 9. Initial model performance on field data using DOE (for 12MW). .....35
Table 10. Model adaption performance using DOE3a on field data for 12MW operation. .....36
Table 11. Model adaption performance using DOE3a (for 14MW field data). .....38
Table 12. HIE model performance metrics (for 14MW) on simulation data. .....39
Table 13. Model adaption performance metrics using DOE3a and 50 samples of 14MW field data. .....40
Table 14. Model adaption performance metrics using DOE3a, 10, 30 and 50 samples of 14MW field data. .....41
Table 15. Data generation process .....53

# Acronyms and Abbreviations

| | |
|---|---|
| AEP | Annual energy production |
| AoA | angle of attack |
| BEM | blade element momentum |
| CBM | Condition based maintenance |
| DEL | damage equivalent load |
| DNV GL | Det Norske Veritas Germanischer Lloyd |
| DOE | design of experiments |
| GWh | gigawatt hours |
| HIE | Health Index Estimation |
| IEC | International Electrotechnical Commission |
| LEE | Leading edge erosion |
| m | meters |
| MAE | mean absolute error |
| MAPE | mean absolute percentage error |
| MW | megawatts |
| NYSERDA | New York State Energy Research and Development Authority |
| O&M | Operation & Maintenance |
| OPEX | Operation expense |
| RUL | Remaining useful life |
| RMSE | root mean square error |
| RSD | relative signal difference |
| TI | turbulence intensity |
| TSR | tip speed ratio |

# Executive Summary

The objective of this project is to develop intelligent blade leading edge health estimation and prognostics models that use turbine operational variables in SCADA as the model inputs. Because of the data sparsity challenge faced in the offshore wind industry, blade health estimation and prognostics models have not been successfully applied to real world offshore turbines. To build such models without a large amount of field data, we proposed to use a high-fidelity Bladed physics model of a 12MW Haliade X prototype turbine to generate simulation data. However, blade leading edge health state per say is not an input signal in Bladed model. Instead, blade roughness, caused by leading edge erosion is represented by airfoil blade section lift and drag aerodynamic coefficients and is a legitimate input to the Bladed model. Therefore, in this project, we aimed to develop the roughness estimation and prognostics models.

This report summarized our accomplishments over the project period from April 2021 to October 2022 on the following four major tasks.

**Task 1** is about generating data using the physical model/simulator. In this task, we developed blade roughness profiles representing the blade leading edge health states, generated the dataset consisting of 11,061 simulation runs and covering a wide range of operating and environmental conditions, and selected, using both univariate-based and multivariate-based qualitative and quantitative data analysis, approximate 20 turbine operation variables that exhibited monotonical trends as the roughness level changes.

**Task 2** is on building health index estimation (HIE) model. In this task, we studied the health index model building technologies by iterating with the variable selection process in Task 1 and developed four HIE models using different machine learning modeling techniques including extreme learning machine (ELM), Random Forest (RF), Linear Regression (LR), and LASSO. By assessing these different HIE models using the four broadly accepted performance metrics, i.e., MAE, RMSE, MAPE and R-squared, we have chosen the ELM model as our HIE model for the rest of our study.

**Task 3** is on calibrating HIE model using the SCADA data. In this task, our focus was on first preparing data from GE's Haliade X prototype as well as defining model calibration strategies, and then calibrating our HIE model for real world adaption. One set of data from the prototype was collected in the first quarter of 2020 that represented different environment condition (Rotterdam of the Netherland) from the simulation. Another set of data from the prototype was collected in the first quarter of 2022 that represented different turbine operation from the design specification (12MW) because the Haliade X

prototype was operated at 14MW in that period. To deal with the environment difference, we developed and implemented a calibration strategy called "calibration via re-simulation" to tune the model for field implementation. To deal with the operation difference, we developed and implemented a calibration strategy called "calibration via model updating" to tune the model for field implementation. By applying the "calibration via re-simulation" to account for environmental conditions, the model performance improved from 79% to 89%, which demonstrates the effectiveness of this procedure and the importance of adapting a standard model to different wind farm locations. By applying the "calibration via model updating" to account for both environment and turbine-to-turbine operational differences (i.e., going from 12MW to 14MW), the model performance improved from 6% to 100%. This calibration method yielded much larger improvements in model adaptation to deal with individual turbine behaviors.

**Task 4** is on building RUL prediction model. In order to plan and schedule condition-based blade maintenance, knowing only the current blade roughness (HIE) is not sufficient. The CBM strategy requires predicting the future roughness growth. Our key focus in this task was to identify and develop a RUL model to forecast the future states of blade roughness. It is worth noting that the RUL forecast in this context is not about remaining useful life to a failure. This forecast specifically determines when the roughness level is likely to cross a predetermined threshold beyond which it is not advisable to operate the turbine without a maintenance action to avoid the AEP loss due to roughness increases.

Challenges in applying the state-of-the-art prognostics models to our application include 1) lacking run-to-failure and historical data; 2) uncertainties from the models, current state, and future state estimations. To tackle these challenges, we selected condition-based prognostics as our method. Condition-based prognostics allows explicit incorporation of various sources of uncertainties. Over the course of project execution, we studied a representative rain erosion lab test data, based on which a baseline RUL model was developed to characterize the increasing trend of roughness over time given a constant rain intensity. Combining the baseline RUL model with the precipitation data in Rotterdam where GE's Haliade X prototype was installed, we simulated a set of data to represent how blade roughness may change over time in the real world. Using the generated data, we developed a roughness RUL model building process and performed RUL model validation to demonstrate the effectiveness of our proposed RUL modeling method.

In summary, our research effort in this project has built and proved a framework for developing HIE and RUL models to realize CBM when historical data and failure data are limited. We believe this framework and its associated techniques are general and can be applied to other failure modes and turbine components of wind turbines.

# 1 Task 1: Generating data using physical model/simulator

## 1.1 Completion simulator preparation

In this milestone, we utilized GE's offshore wind turbine physical model as the framework to simulate how different blade roughness states impact turbine behaviors and performance.

The simulations of the wind turbine were run using Bladed, a commercial software code from DNV GL. Bladed is an aero-elastic, structural dynamics model based on blade element momentum (BEM) theory.

Specifically, within Bladed, we simulated a Haliade X wind turbine. Table 1 is a summary of some main performance parameters of the Haliade X. For more information, please refer the following link. *https://www.ge.com/renewableenergy/wind-energy/offshore-wind/haliade-x-offshore-turbine.*

Table 1. Haliade X descriptive and performance parameters

| Parameter | Specification |
|---|---|
| Rated output (MW) | 12 |
| Rotor diameter (m) | 220 |
| Total height (m) | 248 |
| IEC wind class | 1B |
| Gross AEP (GWhr) | ~68 |

## 1.2 Simulation data generation

A full factorial DOE was assembled to create a wide range of environmental and wind conditions that a real offshore wind turbine might be operated under. Depending on the number of condition variables and levels used, the number of points in the DOE can become very large and unwieldy. Thus, we came up with a suitable space to explore. Figure 1 below is a summary of the condition variables and levels used in the full factorial DOE for a Class 1B turbine. Combining all the variables and associated levels, a total of 648 DOE points resulted from this full-factorial 7 parameter multi-level design.

We used blade roughness levels to parameterize different levels of blade leading edge erosion. In our simulation, we ran 17 roughness levels ranging from 12.5% to 125%.

Figure 1. Parameters and levels used in initial DOE

```
Param1 = [1 1.24];  Air Density in g/cm3
Param2 = [.05 .1834 .25]; Turbulence Intensity (longitudinal)
Param3 = 0.8*Param2; Turbulence Intensity (lateral)
Param4 = 0.5*Param2; Turbulence Intensity (vertical)
Param5 = {'02', '04', '06', '08', '10', '12', '14', '16', '18', '20','22','24'}; Wind Speed in m/s
Param6 = [116 172 175]; Wind file seeds
Param7 = [.1 .14 .2];  Wind vertical shear
```

To simulate the impact of blade roughness on turbine behaviors, we further combined the full-factorial wind and environment condition design described above with 17 roughness state variations to run 11,016 combinations in one large DOE design. That is showed in Figure 2. Each simulation run contains 15,000 datapoints (corresponding to a 10 min time window) with 53 operational variables. The 53 operational variables include both turbine behavioral variables such as blade bending moments and turbine controller estimations such as pitch angle and so on so forth.

Figure 2. Simulation file components


## 1.3 Observability study completion

Given this large number of simulations runs, the key questions to assess roughness observability were –

(1) Do some/any of the turbine behavioral and controller estimation variables or derived features show discernable trends as blade roughness levels are changed under various conditions of DOE?

(2) Are there specific DOE factors and their levels that affect the observability? Correspondingly can we identify a subset of conditions (e.g., wind speed ranges) under which a reliable trend in roughness levels can be identified for prognostics model development.

In this effort, the team conducted a thorough observability study consisting of both univariate-based and multivariate-based qualitative and quantitative data analysis of the simulation data.

After studying the average and individual behaviors of the output variables from the DOE simulation data, we identified about 20 variables that have exhibited more than 5% of changes when the roughness ranges from 35% to 125% [1]. We also found out that effect of roughness on blade operation variables is most evident at lower wind speeds than at higher wind speeds. The effect of roughness gets masked due

to abundance of power available at higher wind speeds. Another noteworthy observation is that that controller estimated variables such as aero torque and tip-speed-ratio etc. show high deviations for larger roughness, which matches our physical intuition since model-based controller estimation has no direct knowledge of roughness, its estimations are based on simulated conditions (as various DOE factors) which remain largely the same across all roughness levels.

Since the only source of stochasticity in simulation is introduced by way of different seeds to simulate wind conditions, the outputs are largely deterministic. To assess robustness of this observability in presence of noise, we simulated noise on several measured variables (i.e., excluding controller estimated parameters). Based on inputs from domain experts random noise varying between 4-7% of the signal amplitude was added to 26 variables, which included torque, blade pitch angle, TSR, and 23 moment variables. From our study, we conclude that observability is preserved despite presence of noise on measurement data.

## 1.4 Discussions of the observations and findings from Task 1

In Task 1, we have completed the preparation of the data simulator and applied the simulator to generate the first batch of simulation data. The observability study determined the operational variables of wind turbines signifies observable difference at different blade roughness levels. Based on the data analysis results obtained, our findings are the following:

1) The resulting changes of the operational variables (sensor measurements, control estimates and derived variables) of wind turbines are algorithmically observable as the blade roughness levels differ.

2) Signature/features calculated from the raw variables have shown good observable trending as the blade roughness increases. Having such signatures with observable trending is of paramount importance for developing blade roughness estimation model in next milestone.

3) The un-measurable factors, e.g., sensor noise and wind shear, of wind turbines have negligible effects on or insensitive to the observability results

# 2 Task 2: Building health index estimation (HIE) model

## 2.1 Completion of feature selection for HIE model

### 2.1.1 Model feature selection

In previous section about observability study, we have determined approximately 20 turbine operational variables present significant changes when the roughness level changes. To build a HIE model to estimate the roughness level, we further performed feature selection to identify a subset of variables (features) that can potentially optimize the model performance while considering the data availability and accessibility in field wind farm operations.

Feature selection or variable selection has been studied in many applications and the corresponding methodologies are chosen based on the problem to be solved. In our case, the HIE model is to characterize the relationship between the roughness level and a number of variables, therefore we would like to identify a transfer function to describe the relationship and the corresponding set of variables that are to be included in the model. We mainly applied three analytics methods in our feature selection task.

The first approach is "Forward stepwise selection in multiple linear regression". A multiple linear regression model assumes the linear relationship between $y$ and each $x$ variable. For each observation $i$ $(i = 1, 2, \dots, n)$:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2\, x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \varepsilon_i,$$

where $\vec{\beta} = (\beta_0, \beta_1, \dots, \beta_k)$ are the coefficients and $\varepsilon_i$ is the error term. The estimated coefficients $\hat{\vec{\beta}} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k)$ are obtained by

$$\hat{\vec{\beta}} = \underset{\vec{\beta}}{\operatorname{argmin}} \sum_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij})^2.$$

The forward stepwise selection algorithm is performed to identify the set of $x$ variables to be included in the multiple linear regression model, which is described in Table 2.

Table 2. Forward stepwise selection algorithm.

| |
|---|
| **Step 1: Start with the initial model without any $x$ variable ($y_i = \beta_0 + \varepsilon_i$) and evaluate its Bayesian Information Criterion (BIC).**<br>**Step 2: Add one more $x$ variable at a time to the previous model (here the initial model) and evaluate its BIC. The resulting model is the one with the largest BIC reduction if any.**<br>**Step 3: Repeat Step 2 until no more $x$ variable can be added to the previous model to reduce the BIC and return the final model with the selected $x$ variables.** |

As an alternative, the least absolute shrinkage and selection operator (LASSO) method regularizes the coefficient estimates, or shrinks the coefficient estimates towards zero. As before, for each observation $i$ $(i = 1, 2, \dots, n)$:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2\, x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \varepsilon_i,$$

where $\vec{\beta} = (\beta_0, \beta_1, \dots, \beta_k)$ are the coefficients and $\varepsilon_i$ is the error term. The LASSO coefficient estimates $\hat{\vec{\beta}} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k)$ are obtained by

$$\hat{\vec{\beta}} = \underset{\vec{\beta}}{\operatorname{argmin}}\{\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij})^2 + l \sum_{j=1}^{k} |\beta_j|\}.$$

The way the coefficients are estimated forces some of the coefficient estimates to be exactly equal to zero when the tuning parameter $l$ increases, hence variable selection is performed. The estimate of $l$ is obtained using the cross-validation technique.

Our third method is Random forests (RF), which is a tree-based ensemble method that aiming to reduce the model variance. A regression tree model doesn't assume the linear relationship between $y$ and $x$ variables. Instead, the high dimensional space of $x$ variables is segmented into a number of small regions each corresponding to a subset of the data. Specifically, a binary splitting is performed at each step using one $x$ variable which leads to the largest error reduction at that step. The binary splitting starts at the top of the tree and successively splits the $x$ variable space each creating two new branches further down on the tree.

A RF model averages a number of noisy but approximately unbiased regression tree models. The RF algorithm is described in Table 3.

Table 3: RF algorithm.

| |
|---|
| Step 1: Draw a bootstrap sample from the training data. |
| Step 2: Grow a regression tree based on the bootstrap sample data, during which at each node a number of randomly selected variables are selected to split the data based on a specified criterion. |
| Step 3: Repeat Step 1-2 a number of times so that multiple regression tree models are developed. |
| Step 4: All the regression trees are ensembled as a "forest" for the model inference. |

There exist various approaches to evaluating the importance or contribution of each $x$ variable during the RF building process. Here we consider two measures of variable importance.

One measure of variable importance is the permutation importance. Notice that each individual tree in the RF is built based on a bootstrap sample of the original data. The data that is not included in the bootstrap sample is called out-of-bag (OOB) data and it can be used for evaluating prediction performance. Permutation importance permutes each $x$ variable's OOB data and compares the resulting OOB prediction error to the original OOB prediction error. The larger the difference in the OOB prediction error, the more important the variable is for the RF model.

Another measure of variable importance is based on a tree-based concept called minimal depth. Notice that variables that split close to the root node of the tree have a strong effect on prediction accuracy and variables that split deeper in the tree have less impact. Minimal depth measures the distance of a variable relative to the root of the tree, and it can be used to assess the prediction impact or importance of the variable. Comparing the minimal depth of each $x$ variable to a threshold derived from the minimal depth distribution filters some noisy variables and therefore can be used for variable selection.

### 2.1.2 Model feature selection observations, findings, and recommendations

The observability study conducted in Task 1 shows that it can be difficult to distinguish roughness levels by using signals at a relatively low wind speed region. Therefore, we only use the data with the normalized wind speed greater than 0.2. Also, some signals at a specific wind speed region can be more likely to distinguish roughness levels than other regions. In this section, we consider two scenarios in

which the normalized wind speed is greater than 0.2, and the normalized wind speed is between 0.4 and 0.5, respectively.

#### *2.1.2.1 Normalized wind speed greater than 0.2*

From the observability study, some signals can interact with air density and TI to distinguish roughness levels. In the multiple linear regression and LASSO, both air density and TI are always included in the model during variable selection. The variable selection results using the methods illustrated in the previous sections are summarized in Table 4, in which the selected variables are marked as "x".

Table 4: Variable selection comparison.

| Variable | Forward stepwise | LASSO | RF minimal depth |
|---|---|---|---|
| Air_Density | x | x | x |
| TI | x | x | x |
| Wind_speed | x | x | x |
| LSS_torque | | x | x |
| Rotor_speed | x | x | x |
| Generator_speed | x | | x |
| Generator_torque | | x | x |
| Generator_air_gap_torque | | x | x |
| Electrical_power | | x | x |
| Reactive_power | x | x | x |
| Blade_pitch_angle | x | | x |
| Hub_wind_direction | x | x | x |
| Rotor_longitudinal_wind_speed | x | | x |
| Rotor_wind_direction | x | | x |
| Blade_Hub_bending moment 1 | x | | x |
| Blade_Hub_bending moment 2 | x | x | x |
| Yaw_bearing_bending moment 1 | | x | x |
| Yaw_bearing_bending moment 2 | x | x | x |
| Nacelle_acceleration 1 | x | x | x |
| Nacelle_acceleration 2 | | x | x |
| Blade_root_bending moment 1 | x | x | x |
| Blade_root_bending moment 2 | x | x | x |
| Blade_root_bending moment 3 | x | x | x |
| Angle of attack 1 | x | x | x |
| Angle of attack 2 | x | x | x |
| Rotating_hub_bending moment 1 | | | |
| Rotating_hub_bending moment 2 | x | x | x |
| Rotating_hub_bending moment 3 | x | x | x |
| Stationary_hub_bending moment 1 | | x | |
| Stationary_hub_bending moment 2 | x | | |
| Stationary_hub_bending moment 3 | x | x | |

| | | | |
|---|---|---|---|
| Tower_bending moment 11 | | x | |
| Tower_bending moment 12 | x | x | x |
| Tower_bending moment 21 | x | x | x |
| Tower_bending moment 22 | x | x | |
| Torque | x | | x |
| Estimated TSR | x | x | x |
| Estimated Power | x | x | |
| Estimated Wind Speed | x | x | x |
| Estimated Wind Speed | x | | x |

### *2.1.2.2 Normalized wind speed between 0.4 and 0.5*

The variable selection results using the methods illustrated in the previous sections are summarized in Table 5, in which the selected variables are marked as "x".

Table 5: Variable selection comparison for selected wind speed range.

| Variable | Forward stepwise | LASSO | RF minimal depth |
|---|---|---|---|
| Air_Density | x | x | x |
| TI | x | x | x |
| Wind_speed | x | x | |
| LSS_torque | | x | x |
| Rotor_speed | | x | x |
| Generator_speed | | | |
| Generator_torque | | | |
| Generator_air_gap_torque | | | |
| Electrical_power | | x | |
| Reactive_power | x | x | |
| Blade_pitch_angle | x | x | x |
| Hub_wind_direction | x | x | x |
| Rotor_longitudinal_wind_speed | x | | |
| Rotor_wind_direction | x | x | |
| Blade_Hub_bending moment 1 | x | x | |
| Blade_Hub_bending moment 2 | x | x | |
| Yaw_bearing_bending moment 1 | | x | x |
| Yaw_bearing_bending moment 2 | x | | |
| Nacelle_acceleration 1 | | x | |
| Nacelle_acceleration 2 | | x | |
| Blade_1_root_bending moment 1 | x | x | |
| Blade_1_root_bending moment 2 | x | x | |
| Blade_1_root_bending moment 3 | x | x | |
| Angle of attack 1 | x | x | x |
| Angle of attack 2 | x | x | |
| Rotating_hub_bending moment 1 | | | |
| Rotating_hub_bending moment 2 | x | x | x |
| Rotating_hub_bending moment 3 | | x | |

| | | | |
|---|---|---|---|
| Stationary_hub_bending moment 1 | | | |
| Stationary_hub_bending moment 2 | x | x | |
| Stationary_hub_bending moment 3 | x | x | |
| Tower_bending moment 11 | x | x | |
| Tower_bending moment 12 | | | |
| Tower_bending moment 21 | x | x | |
| Tower_bending moment 22 | x | | |
| Torque | x | | |
| Estimated TSR | x | | x |
| Estimated Power | | x | |
| Estimated Wind Speed | x | x | x |
| Estimated Rptor Speed | x | x | |

## 2.2 Health index estimation (HIE) modeling

A health state estimation or health index estimation (HIE) model takes turbine's operational measurements (e.g., SCADA data and control estimates) at a given time as the inputs and outputs the blade roughness level, a health index, at that time. The estimated blade roughness levels or health indices are used not only for accurately monitoring the current condition of the blade, but also for propagating the condition state to future operating state for remaining useful life (RUL) estimation, which ultimately enables for effectively and confidently exercising a CBM strategy on offshore wind turbine blades.

### 2.2.1 HIE modeling strategies and techniques

The goal of HIE modeling is to develop a ML-based predictive model that estimates the roughness status or levels of a wind turbine based on its operational measurements (sensor measurements and control estimates). A wind turbine is a complex, dynamic system that operated in stochastic conditions; and thus the operational measurements collected from the turbine system are certainly affected by many contributing factors, including wind speed, wind directions, air density, etc., in addition to the blade roughness changes. The challenges associated with blade roughness estimation modeling come from the fact that not all these contributing factors are measured and available in real-world wind turbine applications. Hence, minimizing the effect of these non-measurable factors to blade roughness status becomes the key to the success of our HIE modeling. Our observability study, sensitivity analysis completed in Phase I and variable/feature selection are a part of our efforts towards addressing the challenges and have provided us an initial set of important variables. For HIE modeling in this study, we start our HIE model inputs with the smallest number of variables that are most widely available in real-world turbines, and more variables are added as the HIE model inputs only when these added variables significantly improve the model prediction performance.

From machine learning perspectives, blade roughness or health index estimation (HIE) is a regression problem, where the independent variables or inputs are operational measurements, e.g., measured wind speed and measured pitch angle, and the dependent (or target) variable is simply the blade roughness level or health index. In literature there are many regression methods available, ranging from traditional statistical approaches, e.g., linear regression, to most advanced machine learning methods. For our HIE modeling we adopted the Extreme Learning Machine (ELM) [1] as the primary HIE modeling technique, due to its unique advantages over other modeling techniques (see below). We also explored other machine learning models, e.g., random forest, linear regression, as a part of our compassion study.

Extreme learning machine (ELM) is a special type of feed-forward neural networks [2]. ELM and random vector functional links (RVFL) and others forma special group of neural networks called randomized neural networks. In traditional feed-forward neural networks, network training, i.e., finding all connection weights and bias, is performed through an iterative optimization. ELM, on the other hand, follows a non-iterative learning approach, that is, the connections between input and hidden neurons are randomly generated and fixed; and training an ELM becomes finding connections between hidden and output neurons only, which is simply a linear least squares problem that can be solved analytically. Because of such special design of the network, ELM training is significantly fast. Moreover, ELM has approved to be efficient and effective for many applications (both classification and regression) [3].

Consider a dataset, $\{(\boldsymbol{x}_i, y_i)\}_{i=1}^{N}$, $\boldsymbol{x_i} \in \mathbb{R}^{\boldsymbol{d}}, y_i \in \mathbb{R}^k$ , and a network with *L* hidden neurons. Then the output of the network is expressed as:

$$\boldsymbol{f}(\boldsymbol{x}) = \sum_{i=1}^{L} \boldsymbol{\beta_i} h_i(\boldsymbol{x}) = \boldsymbol{h}(\boldsymbol{x})\boldsymbol{\beta} \tag{1}$$

Where $h_i(\boldsymbol{x}) = G(\boldsymbol{w}_i, b_i, \boldsymbol{x})$, $\boldsymbol{w}_i \in \Re^d, b_i \in \Re^1$, is the output of $i^{th}$ hidden neuron with respect to the input $\boldsymbol{x}$; $G(w, b, x)$ is a nonlinear piecewise continuous function satisfying ELM universal approximation capability theorems [5]; $\boldsymbol{\beta}_i$ is the output weight vector between $i^{th}$ hidden neuron to the $k \geq 1$ output nodes. $\boldsymbol{h}(\boldsymbol{x}) = [h_1(\boldsymbol{x}), \ldots, h_L(\boldsymbol{x})]$ is a random feature map mapping the data from d-dimensional input space to the *L*-dimensional random feature space (or the ELM feature space).

For the equality optimization constraints-based ELM, the unknown parameter, $\boldsymbol{\beta}$, is found through the following optimization:

$$\text{Minimize: } \boldsymbol{L}_p = \frac{1}{2}\|\boldsymbol{\beta}\|^2 + \frac{1}{2}C\sum_{i=1}^{N}\|\boldsymbol{\xi}_i\|^2$$

$$\text{Subject to: } \boldsymbol{h}(\boldsymbol{x}_i)\boldsymbol{\beta} = \boldsymbol{y}_i^T - \boldsymbol{\xi}_i^T, i = 1, \dots, N \tag{2}$$

Where $\boldsymbol{\xi}_i = [\xi_{i,1}, \dots, \xi_{i,k}]^T$ is the training error vector of the *k* output nodes with respect to the training sample $\boldsymbol{x}_i$ and the constant *C* controls the tradeoff between the output weights and the training error.

The equivalent dual optimization objective function of Eq. 2 is:

$$\boldsymbol{L}_d = \frac{1}{2}\|\boldsymbol{\beta}\|^2 + \frac{1}{2}C\sum_{i=1}^{N}\|\boldsymbol{\xi}_i\|^2 - \sum_{i=1}^{N}\sum_{j=1}^{k}\alpha_{i,j}(\boldsymbol{h}(\boldsymbol{x_i})\boldsymbol{\beta_j} - y_{i,j} + \xi_{i,j}) \tag{3}$$

Based on the Karush-Kuhn-Tucker (KKT) condition, we can have the solutions for the ELM output function $\boldsymbol{f}(\boldsymbol{x})$ for non-kernel and kernel cases, respectively, as follows.

*Non-kernel case*

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{h}(\boldsymbol{x})\boldsymbol{\beta} = \boldsymbol{h}(\boldsymbol{x})\boldsymbol{H}^T\left(\frac{\boldsymbol{I}}{\boldsymbol{C}} + \boldsymbol{H}\boldsymbol{H}^{\boldsymbol{T}}\right)^{-1}\boldsymbol{Y} \tag{4}$$

where $\boldsymbol{H}$ is the hidden layer output matrix.

$$\boldsymbol{H} = \begin{bmatrix}\boldsymbol{h}(\boldsymbol{x_1})\\ \vdots \\ \boldsymbol{h}(\boldsymbol{x_N})\end{bmatrix} = \begin{bmatrix}h_1(\boldsymbol{x}_1) & \dots & h_L(\boldsymbol{x}_1)\\ \vdots & \vdots & \vdots \\ h_1(\boldsymbol{x}_N) & \dots & h_L(\boldsymbol{x}_N)\end{bmatrix} \tag{5}$$

*Kernel case*

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{h}(\boldsymbol{x})\boldsymbol{H}^T\left(\frac{\boldsymbol{I}}{C} + \boldsymbol{H}\boldsymbol{H}^{\boldsymbol{T}}\right)^{-1}\boldsymbol{Y} = \begin{bmatrix}\boldsymbol{K}(\boldsymbol{x}, \boldsymbol{x_1})\\ \vdots \\ \boldsymbol{K}(\boldsymbol{x}, \boldsymbol{x_N})\end{bmatrix}^{\boldsymbol{T}}\left(\frac{\boldsymbol{I}}{C} + \boldsymbol{\Omega}_{\boldsymbol{ELM}}\right)^{-1}\boldsymbol{Y} \tag{6}$$

Where $\boldsymbol{\Omega}_{\boldsymbol{ELM}} = \boldsymbol{H}\boldsymbol{H}^{\boldsymbol{T}}$, *and* $\Omega_{ELM\ i,j} = \boldsymbol{h}(\boldsymbol{x_i}) \cdot \boldsymbol{h}(\boldsymbol{x_j}) = \boldsymbol{K}(\boldsymbol{x_i}, \boldsymbol{x_j})$.

### 2.2.2 Model performance metrics and evaluation methods

The performance of a regression model is often measured by prediction errors, difference between the predicted and the true target values. For HIE modeling, in this report we utilize four different performance metrics, for assessing our HIE models' performance. They are the root mean squared error (RMSE), the mean absolute error (MAE), the mean absolute percentage error (MAPE), and the R-square values. Given N data samples, the target and the predicted values are $y_i$ and $\hat{y}_i$, respectively, the four error metrics are defined as follows.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{7}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{8}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{(y_i - \hat{y}_i)}{y_i}\right|^2 \tag{9}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{10}$$

In terms of performance evaluation method, we adopt the 60/40 hold-out cross-validation method. That is, the entire data is randomly shuffled and then split into two disjoint subsets, 60% and 40%, while the 60% of it is used for training the HIE model and the rest of data (40%) is used for testing the model's generalization capability. We use the performance metrics on the testing dataset for model validation and model performance comparison.

### 2.2.3 HIE modeling results

#### *2.2.3.1 HIE model details*

*Model inputs.* Based on the sensor measurements availability of a typical real-world turbine application, we have chosen seven variables as shown in Table 6 as our HIE model inputs. These seven variables are very basic measurements that are available for almost all wind turbine operations/applications. These seven variables are also a subset of variables selected in our variable selection study conducted in the previous quarter.

Table 6: Input variables of our HIE model.

| Model Inputs | Variable Names |
|---|---|
| X1 | Cup_anemometer_wind_speed |
| X2 | TI |
| X3 | Air density |
| X4 | Blade pitch angle |
| **X5** | Controller estimated wind speed |
| **X6** | Blade moments |
| X7 | Tip-speed ratio |

And thus our HIE model is then defined as:

$$roughness = f(X_1, X_2, \ldots X_7)$$

where, *f* is the ELM model to be built.

*Input variable scaling/normalization*. To ensure a proper training and a good performance of our HIE model, each of the input variables is scaled to [0, 1] using range scaling.

*Model hyperparameters*. The ELM model has two hyperparameters, the number of hidden neurons and the regularization factor C (see Eq. 2). These two hyperparameters are often determined via cross-validation. In this study, the final values of these two hyperparameters are 300 and 2^16, respectively.

### *2.2.3.2 Model performance summary*

We used 60% of the data samples for training and the remaining 40% data samples for testing our ELM model. Figure 3 shows the scatter plots between the true and the predicted rough levels for both the training and the testing set, respectively. The distributions (histograms) of the predictions for each of the six roughness levels are shown in Figure 4. From these figures one can see that the ELM model predicts the blade roughness levels reasonably well, especially at the higher levels, e.g., 50% and above, of blade roughness. One can also see that the ELM model has some difficulty differentiating clean (12.5% roughness) and small (35%) roughness level, which makes physical sense and is well understandable.

To quantitatively assess the ELM model performance, we calculated the four performance metrics defined in Section Model performance metrics and evaluation methods2.2.2. For the training dataset, the MAE, the RMSE, the MAPE and the R-squared were 3.82, 4.86, 15.06, and 0.98, respectively, while for the testing dataset, they were 3.88, 4.89, 13.72 and 0.98, respectively.

We have also explored other different modeling techniques for our HIE model. Since in the earlier efforts for feature selection (M2.1) we have utilized three modeling techniques, e.g., random forest (RF), LASSO and linear regression (LR), for variable selection. We thus also built three additional HIE models using these three modeling techniques.

Table 7 shows the comparison of prediction performance between our ELM and the three additional models, in terms of the four performance metrics. From the table one can clearly see the ELM model outperformed other models significantly, which further validates the superiority of our ELM-based HIE model for predicting blade roughness levels.

Figure 3: Scatter plot between the true and the predicted roughness levels.

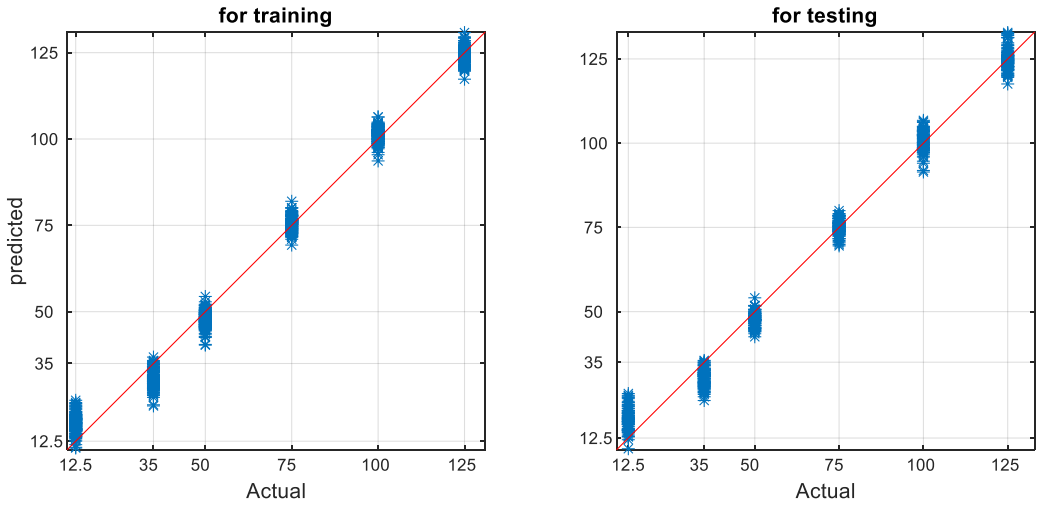



Figure 4: Histograms of the predicted roughness levels.

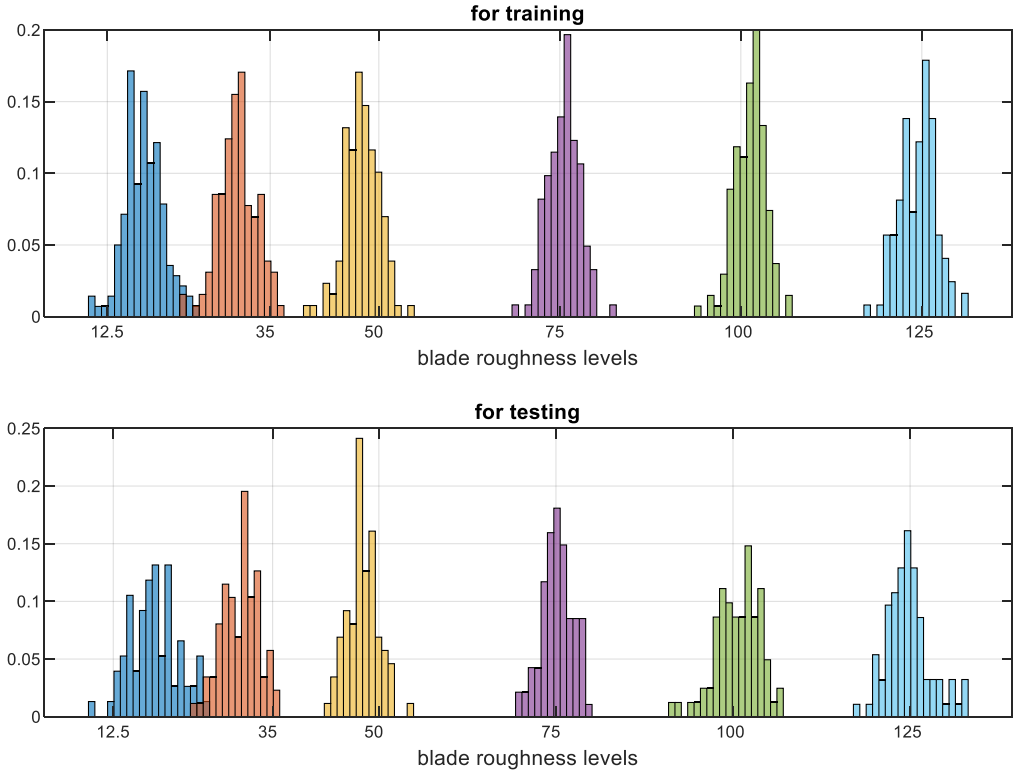


Table 7: HIE model performance comparison.

| Perf. metrics | Training | | | | Testing | | | |
|---|---|---|---|---|---|---|---|---|
| | ELM | RF | LASSO | LR | ELM | RF | LASSO | LR |
| MAE | 3.82 | 3.19 | 7.52 | 7.41 | 3.88 | 8.55 | 7.56 | 7.63 |
| RMSE | 4.86 | 4.11 | 9.53 | 9.27 | 4.89 | 11.29 | 9.49 | 9.49 |
| MAPE | 15.06 | 11.33 | 25.97 | 24.25 | 13.72 | 31.8 | 26.47 | 25.06 |
| R-squared | 0.98 | 0.99 | 0.94 | 0.94 | 0.98 | 0.92 | 0.94 | 0.94 |

## 2.3 Discussions of the observations and findings from Task 2

In Task 2, our focus has been on the HIE modeling. We developed our blade roughness estimation model based on the simulated data. Realizing that many factors (or contributing factors) affect blade roughness and that not all these contributing factors are measurable, during our estimation model building, we made a great effort on minimizing the effect of these non-measurable factors by limiting the number of input variables to the smallest possible and ensuring that these input variables are widely available in typical real-world wind turbine operations. Based on the simulation data, we have demonstrated that our HIE model developed performed reasonably well in terms of the four-performance metrics and our HIE model developed significantly outperformed other estimation models we considered.

# 3 Task 3: Calibrating HIE model using the SCADA data

## 3.1 Completion of filed data preparation and calibration strategy

### 3.1.1 Motivation for calibration

The early deliverables reported that the simulation data was generated based on the specification of wind turbine design parameters, controller configuration, and environmental inputs. While a wide range of simulation scenarios can be generated for model building, we expect unseen conditions to appear in fielded systems, which may not necessarily be covered by the range of inputs in simulation data used for model training. Therefore, the relationship between the blade roughness and other variables can be different if any of these inputs change beyond the pre-defined range. In general, if any input to the simulation based on which an HIE model is built is significantly different from what is observed in the field, a model calibration process may be needed to retune the HIE model so that it better predicts blade roughness state in the field.

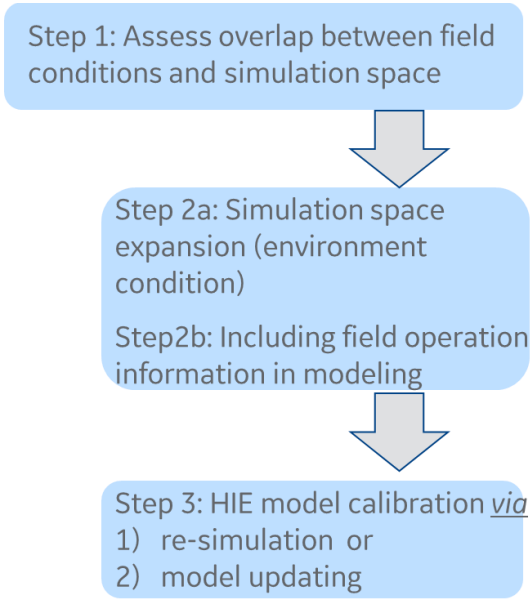
For instance, to validate the HIE model developed using simulation data against the field test, we will apply the developed HIE model to the field data and study how well the model performs. Field conditions are more dynamic and complex than those simulated through a model. This often results in a poorer overall performance on field data as compared to validation studies based on simulation data alone. In such cases, a model calibration step is needed to calibrate and improve the HIE model so that it can be adapted to the field data for an accurate estimation and monitoring of the blade roughness level in real time. As a specific example, in a new DOE simulation design, we obtained wind turbine design parameters and control configurations from the fielded turbine and leveraged these same inputs in the simulation to mimic wind turbine deployment as closely as possible. This allows us to isolate the effects of changing environmental inputs and hence characterize potential differences between the simulation and the fielded system due to environmental factors alone.

Since the installation in 2019, the Haliade X turbine has been operated to generate different rated powers (12MW, 13MW and 14MW). Key control variable such as blade pitch angle varies significantly as a result. A calibration strategy for this scenario needs to be developed to deal with both environment and operation differences.

### 3.1.2 Model calibration strategy

In both calibration scenarios, we follow three steps showed in Figure 5 to show that an HIE model can be adapted to the field data by (Step 1) assessing the overlap between the simulation input space and actual field conditions, (Step 2a) expanding the simulation input space as needed, or (Step2b) incorporating turbine control specific variables into model updating, and (Step 3) updating and validating the HIE model using additional simulation data accordingly in Scenario 1 or using additional field data in Scenario 2.

Figure 5. Calibration steps.




## 3.2 Completion of HIE model calibration and validation

### 3.2.1 Apply model calibration to field data

The calibration strategy for Scenario 1 was proposed based upon the data analysis insights obtained from GE's Haliade X offshore wind turbine prototype operated at 12MW capacity in the first quarter of 2020. Specifically, this so called "calibration via re-simulation" strategy deals with different environmental condition between Haliade X prototype in Rotterdam and the simulation.

We later got access to a more recent dataset from Haliade X when it was operated at 14MW rated power in January of 2022. We named this case as Scenario 2. The data analysis insights from the new dataset

necessitated a new calibration strategy called "calibration via model updating". Specifically, this strategy provides a calibration process for turbine operation differences, where it is necessary to exercise simulation to generate more data and also update model.

Details about Scenario 1 - the "calibration via re-simulation" are described in Section 3.2.2 and details about Scenario 2 are presented in Section 3.2.3.

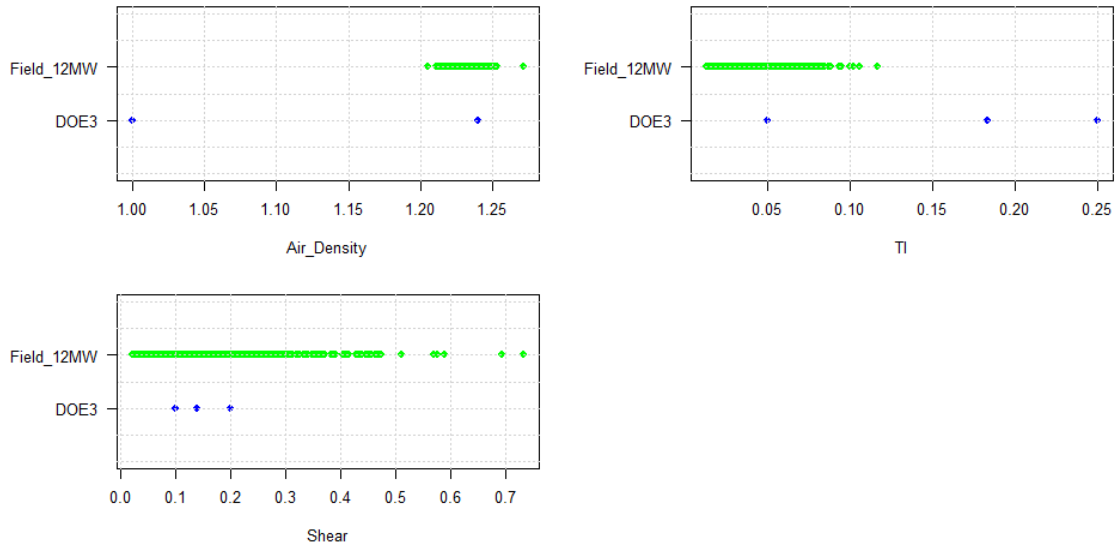
### 3.2.2 Scenario 1 – "Calibration via re-simulation"

In this section, we first describe the field data analysis that prompts the calibration Scenario 1, where we observed previously unseen environmental conditions that the DOE simulation didn't include, or the included data points were sparse. We then implement our calibration method and show the improved model performance through field data validation with GE's Haliade X operated at 12MW capacity. Our simulation was setup with design settings for 12MW operation, and we expect operating environment characteristics to be the key factor for any differences. Specifically, we need to make sure the range of environmental factors used in simulation DOE matches the range observed in the field.

#### *3.2.2.1 Assessing overlap between field conditions and simulation input space*

As Step 1 in the model calibration strategy, understanding the real-world data and individual turbine behavior is critical to adapt the HIE model built with simulation data to real-world so the model can be eventually implementable on real turbines and trusted for estimating and predicting turbine health conditions. We received help from GE's Offshore business team in accessing the field data from the Haliade-X prototype. A set of field data was collected from Haliade-X prototype that was operated at the rated power of 12MW between February 2020 and April 2020. Figure 6 summarizes and compares the air density, TI, and shear for the DOE [4] simulations and the field data. As can be seen the two datasets cover different ranges of the three environmental parameters. Specifically, for instance, a larger spread is seen in air density around 1.25. TI also shows a larger spread but also includes ranges that are not within simulated TI ranges. Likewise, shear values in field data span a much wider range than what was originally simulated. As the relationship between the blade roughness and other variables is expected to be non-linearly different under different environmental conditions, we cannot expect a model trained on narrower range to perform reasonably on full range of field conditions. Therefore, there was a need to calibrate the model by adjusting the simulation input space and cover the field conditions that were not previously included in the DOE simulation data.

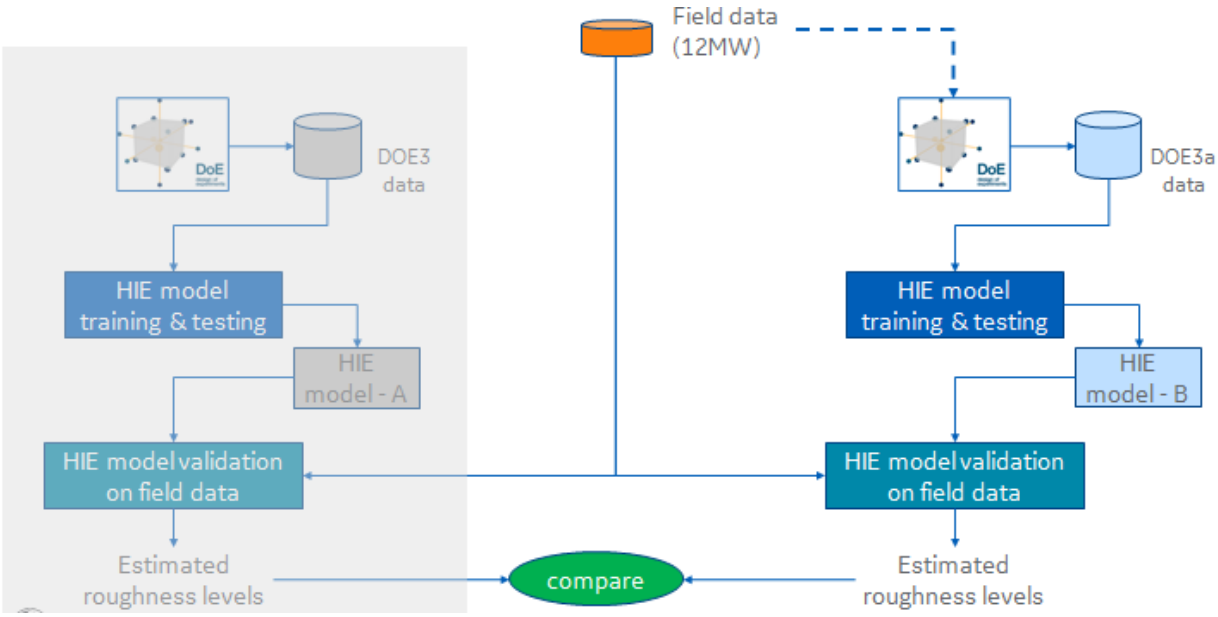
Figure 6. Environmental condition comparison for DOE and 12MW field data.




### *3.2.2.2 Calibration process – "calibration via re-simulation"*

For Scenario 1, where the simulation and the field data (12MW) have different environmental conditions, e.g., TI, air density and shear, we employ "calibration via re-simulation" approach shown in Figure 7, that is, simply to "customize" the design space of the DOE simulation to better match the environmental conditions of the field data (12MW). Then the HIE model is re-trained based on the simulation data with the customized design space. We demonstrate the effectiveness of our "calibration via re-simulation" by comparing the estimated roughness accuracies on the field data (12MW) between the two HIE models, HIE model built on the original simulation data (DOE) and the HIE model built on the simulation data with the customized design space, respectively. Note that the two HIE models have identical model structure, model hyperparameters, and input variables.

Figure 7 Calibration via re-simulation.




### *3.2.2.3 Simulation space expansion*

To adjust the simulation input space (Step 2a) to better match the field conditions, we generated another DOE simulation data (DOE3a) with a full factorial design. A 3D view of all three environmental conditions for the new DOE(DOE3a) and previous DOE(DOE3) is illustrated in Figure 8. The comparison of air density, TI, and shear for the original DOE, and updated DOE and field data is

displayed in Figure 9, which shows that the environmental conditions for the new simulation data are more consistent with those of the field data from the Haliade X prototype

Figure 8. 3D view of the environmental condition for DOE and DOE3a.

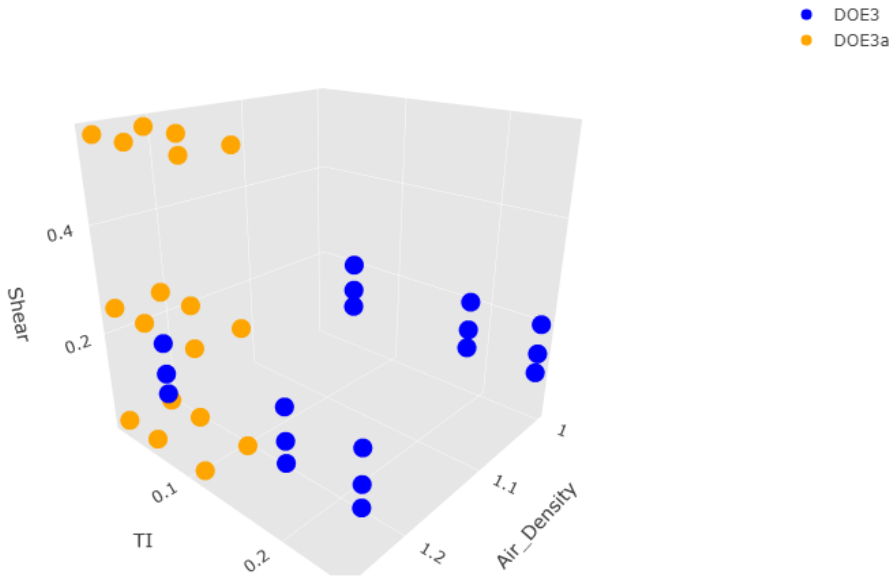
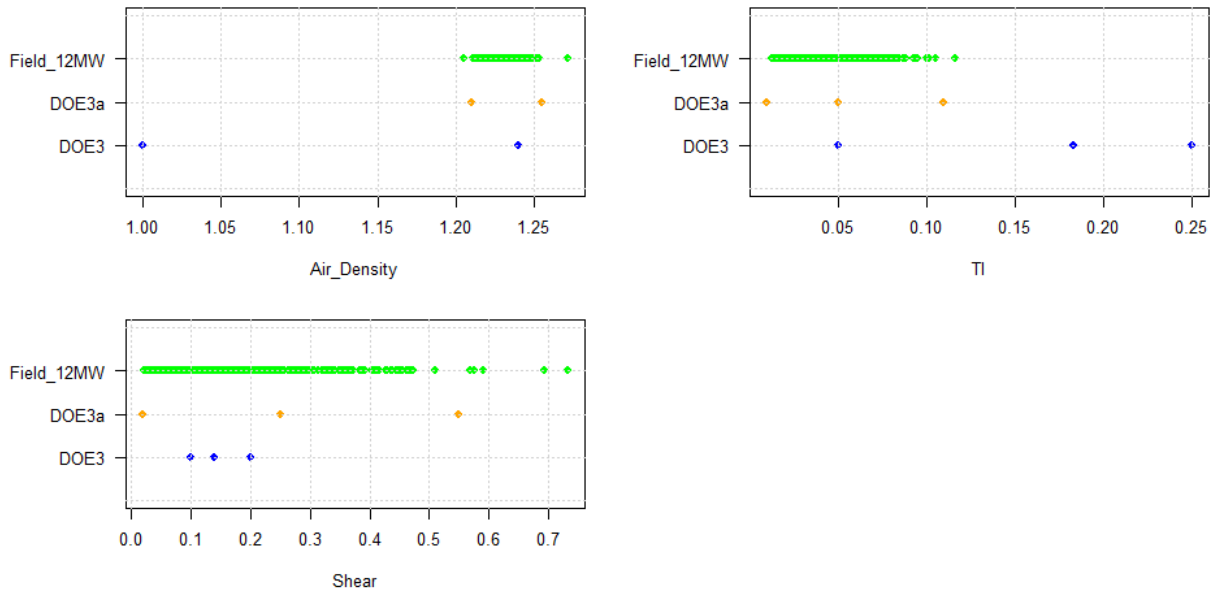



Figure 9. Environmental condition comparison for DOE, DOE3a and 12MW field data.




### *3.2.2.4 HIE model adaption to the 12MW field data*

As Step 3 in Scenario 1, the newly simulated dataset DOE3a and the 12MW field data from the Haliade X

prototype are now incorporated in the model calibration.

To use the updated DOE3a data to develop an HIE model (Step 3) that can be adapted to the field data, we will need to assess the availability and associated data quality for the variables available from the field. It was determined that the available field data included only a subset of the variables compared to the full set that was generated in our simulation.

We developed the HIE model based on the DOE3a data using the Extreme Learning Machine (ELM) technique as discussed in Task 2. Figure 10 shows the scatter plots between the true and the predicted roughness levels for both the training and the testing set, respectively. The distributions (histograms) of the predictions for each of the six roughness levels are shown in Figure 11. Both figures show that the ELM model predicts the blade roughness levels reasonably well. The model performance metrics are also shown in

Table 8.

Figure 10. Scatter plot between the true and the predicted roughness levels on simulated data.

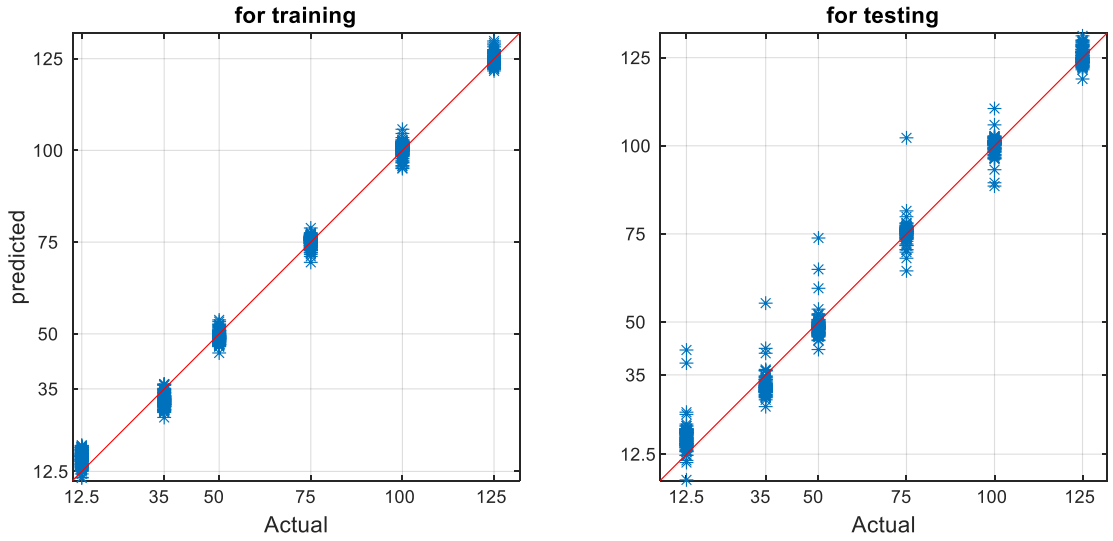


Figure 11. Histograms of the predicted roughness levels (for 12MW simulation data).

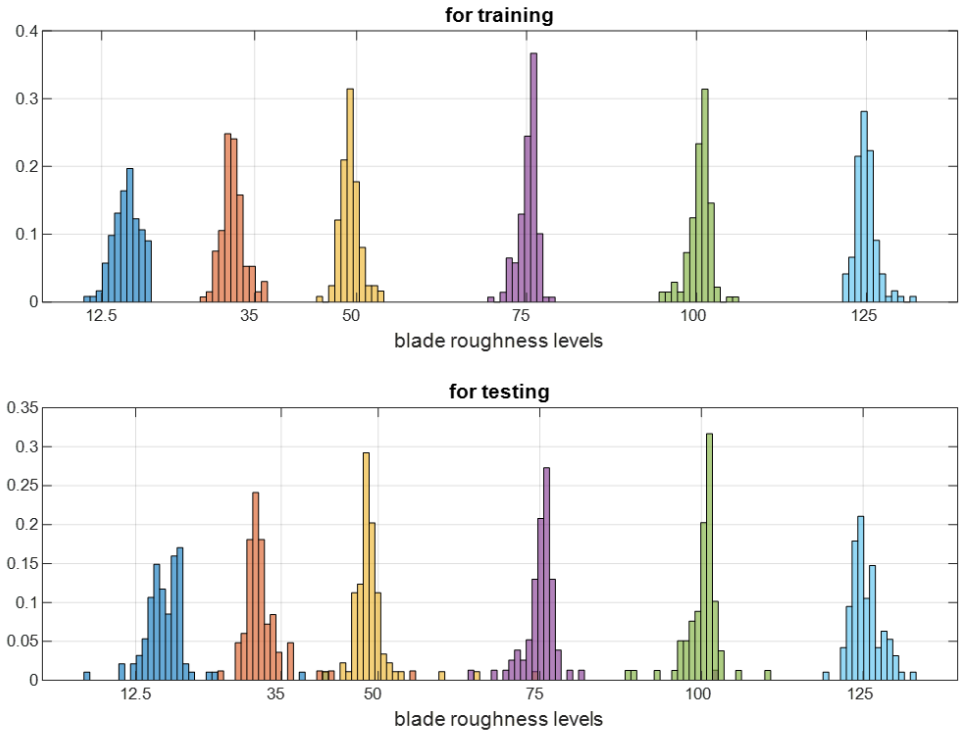



Table 8. HIE model performance on simulated data (train and test).

| | MAE | RMSE | MAPE | R-squared |
|---|---|---|---|---|
| **Training** | 1.935 | 2.500 | 7.110 | 0.996 |
| **Testing** | 2.751 | 4.027 | 10.386 | 0.990 |

Next for validation, the above developed HIE model was applied to the 12MW field data as is and model performance was assessed. As the field data was collected from Haliade-X only a few months after its installation, we assume that the actual blade roughness status is clean.

As a baseline, Figure 12 and Table 9 show the model prediction results for the 348 field data points (10 minute timeseries for each point) using the HIE model built upon the DOE data, in which 275 (79%) data points are classified as clean (less than 50% rough), but a large number of points (21%) were predicted by the model to be very rough (greater than 50% rough), which are likely candidates for false positives if our assumption of clean blade holds. To demonstrate calibration strategy through input space coverage

assessment and subsequently simulation space expansion we further improve the model with additional data from DOE3a. As shown in Figure 13 and Table 9, 308 data points are now predicted as "clean" (50% or less rough), and the prediction accuracy has improved to 89%.

Figure 12. Initial model performance on field data using DOE (for 12MW).

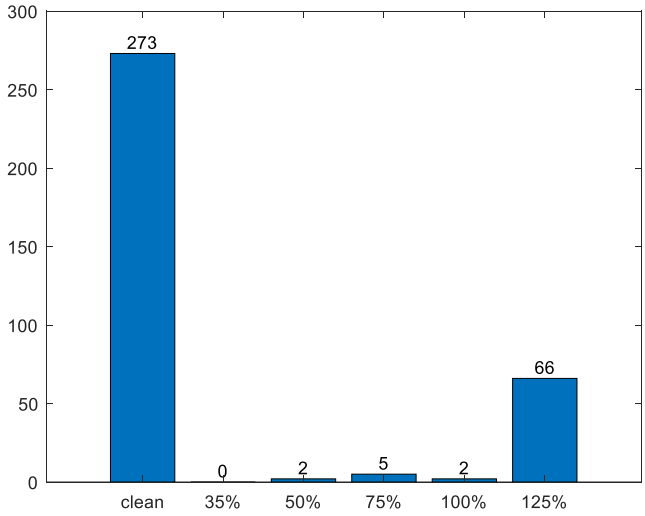



Table 9. Initial model performance on field data using DOE (for 12MW).

| Accurate prediction | Inaccurate prediction |
|---|---|
| 79% | 21% |

This further verifies our model calibration strategy for Scenario 1 that adjusting simulation input space can improve the model adaption performance. In an operational setting, data from specific site can be analyzed and used for generating additional simulation data from new conditions to adapt/calibrate the HIE model to new conditions.

Figure 13. Model adaption performance using DOE3a (for 12MW) on field data.

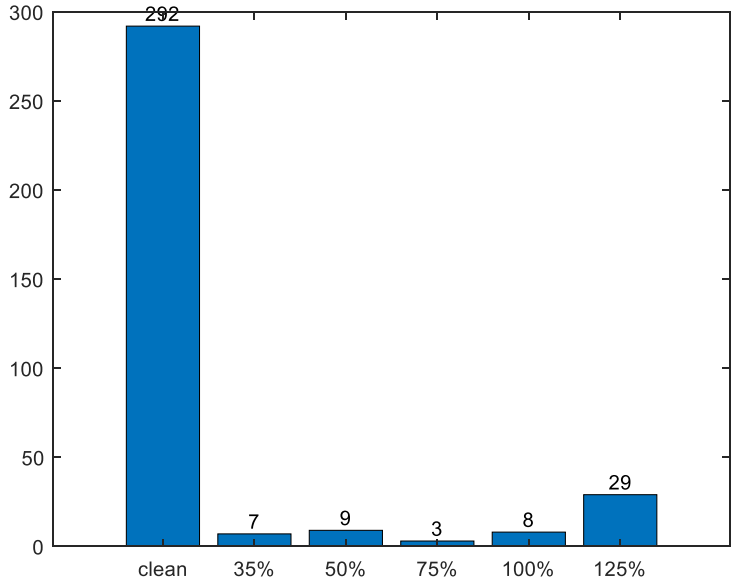



Table 10. Model adaption performance using DOE3a on field data for 12MW operation.

| Accurate prediction | Inaccurate prediction |
|---|---|
| 89% | 11% |

### 3.2.3 Scenario 2 – Calibration via model updating

In this section, we describe the field data analysis that manifests calibration scenario 2, where we not only observed unseen environmental conditions, but also the control and operation variables that the DOE simulation didn't include. This case was named as Scenario 2. We then implement our calibration method, "Calibration via model updating", and show the improved model performance through field data validation with GE's Haliade X operated at a higher 14MW rated power.

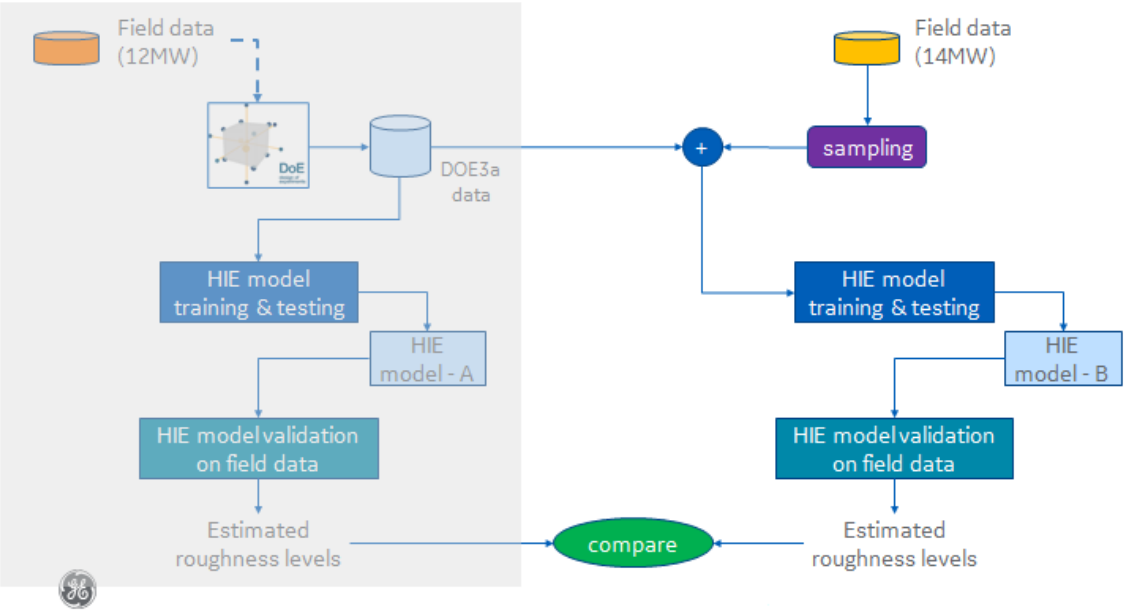
#### *3.2.3.1 Assessing field conditions and turbine control variables*

A second set of field data was recently collected from the Haliade-X prototype that was operated at a higher rated power of 14MW in January 2022. Figure 12 summarizes all the available variables, in which the ranges of some turbine controller related variables are very different from those in the 12MW field data due to the different turbine operational mode.

### *3.2.3.2 Calibration process - "calibration via model updating"*

The data comparison in the previous subsection clearly shows significant environmental condition differences between the simulation (DOE3a) and the field data (14MW). Consequently, the HIE model trained on the simulation data will not work well for estimating roughness levels, when applied to the field data, since machine learning models inherently are not well generalizable to out-of-distribution samples (note that the 14MW field data has mostly out-of-distribution samples). To ensure our HIE model works on both the simulation and the field data as well, we perform model calibration via model updating as showed in Figure 14. Specifically, we update our HIE model with a small number of samples randomly drawn from the field data. The effectiveness of this approach is presented next by comparing the estimated roughness accuracies on the field data between the two HIE models, HIE model built on the simulation data only (without calibration) and the HIE model built on both simulation and sampled field data (with calibration), respectively. Note that the two HIE models have identical model structure, model hyperparameters, and input variables.

Figure 14. Calibration via model updating.




### *3.2.3.3 HIE model updating to the 14MW field data*

We first developed an HIE model using the DOE3a data and the variables that are available in both the DOE3a and 14MW field data. The developed HIE model was then applied to the 14MW field data. The model prediction results are shown in Figure 15 and Table 11, in which only 13 (6%) out of 222 data points are correctly classified as clean (less than 50% rough).

Figure 15. Model adaption using DOE3a performance on 14MW field data.

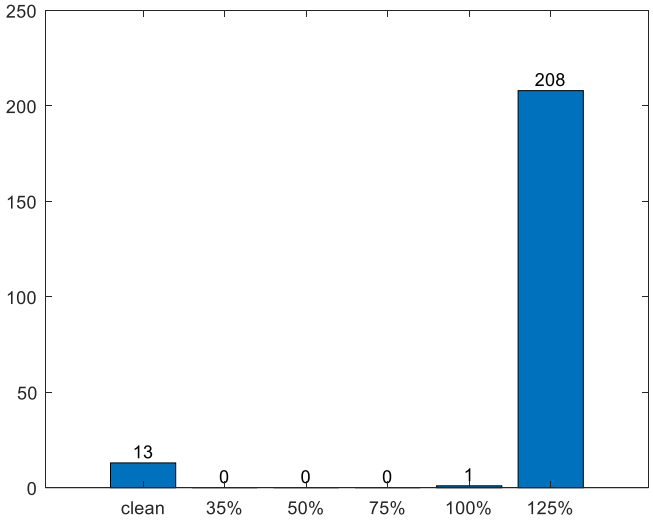



Table 11. Model adaption performance using DOE3a (for 14MW field data).

| Accurate prediction | Inaccurate prediction |
|---|---|
| 6% | 94% |

To calibrate the HIE model for better model adaption in this scenario, we combined the DOE3a simulation data with a subset of the 14MW field data for model training. With only 50 samples from the 14MW field data selected, an HIE model was developed with good training performance as shown in Figure 16, Figure 17 and Table 12.

Figure 16. Scatter plot between the true and the predicted roughness levels on simulation data (for 14MW).

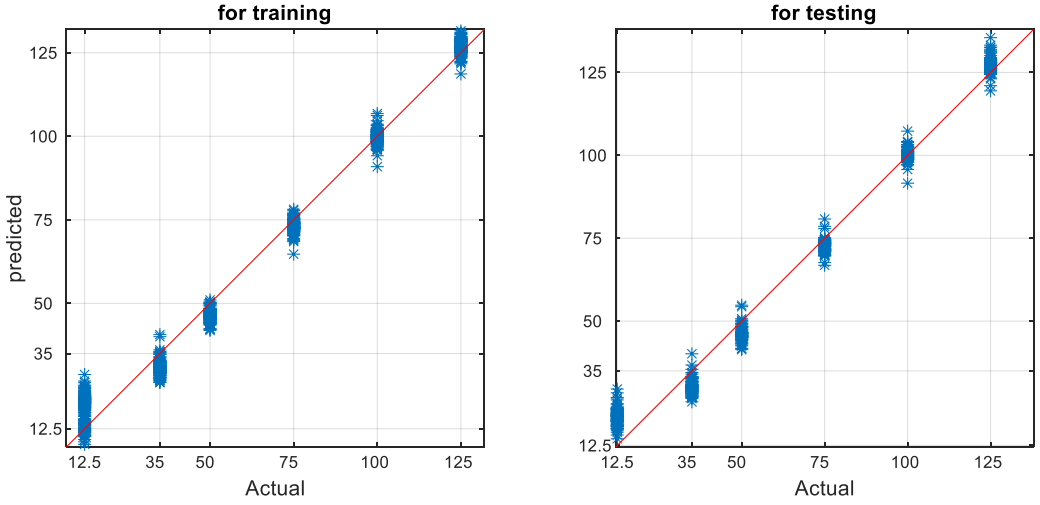



Figure 17. Histograms of the predicted roughness levels (for 14MW) on simulated data.

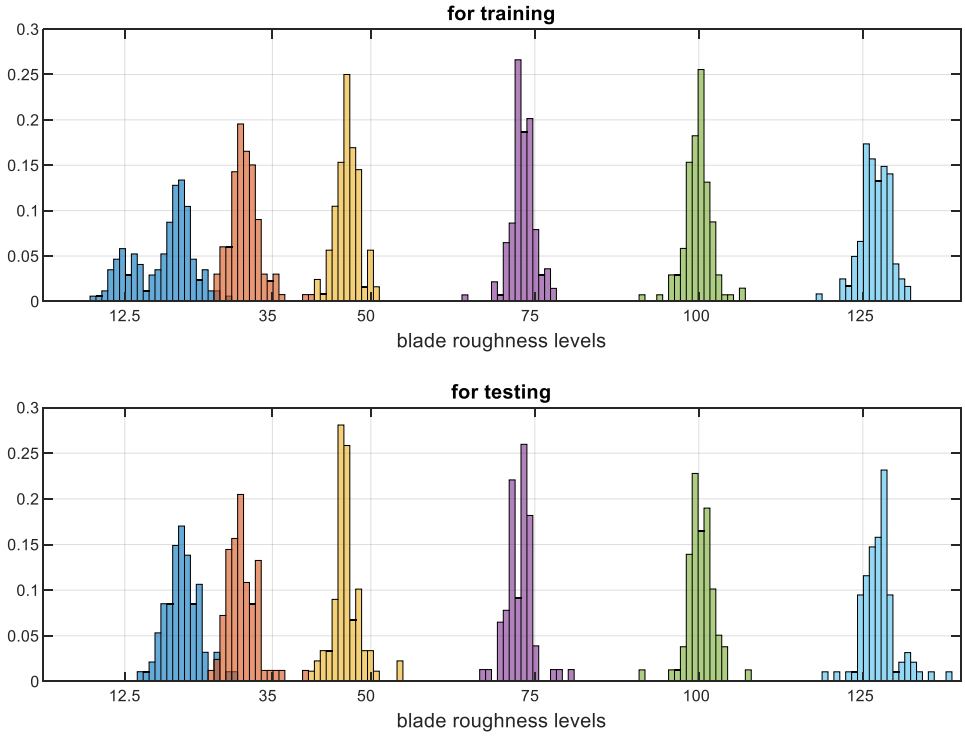



Table 12. HIE model performance metrics (for 14MW) on simulation data.

| | MAE | RMSE | MAPE | R-squared |
|---|---|---|---|---|
| **Training** | 3.514 | 4.516 | 14.726 | 0.986 |
| **Testing** | 4.179 | 5.205 | 17.408 | 0.982 |

The above developed HIE model is then applied to the 14MW field data, and the model prediction results are shown in Figure 18 and Table 13 in which all the 222 (100%) data points are correctly classified as clean. This verifies our model calibration strategy that leveraging field data in the HIE model development can also improve the model adaption performance.

Figure 18. Model adaption performance using DOE3a and 50 samples of 14MW field data.

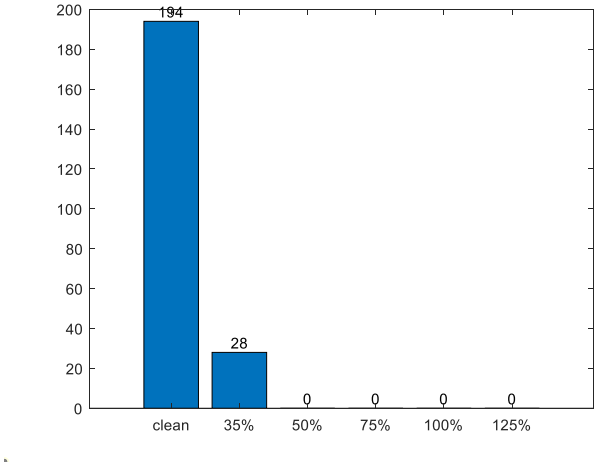



Table 13. Model adaption performance metrics using DOE3a and 50 samples of 14MW field data.

| Accurate prediction | Inaccurate prediction |
|---|---|
| 100% | 0% |

To investigate the requirement on the number of field data sample for model adaption, we also conducted a sensitivity analysis by updating the model with 10 and 30 field samples. The comparison results including using 50 samples are shown in Figure 19 and Table 14. From this investigation, we conclude that the model updating with more field data yields better performance.

Figure 19. Model adaption performance using DOE3a and 50 samples of 14MW field data.

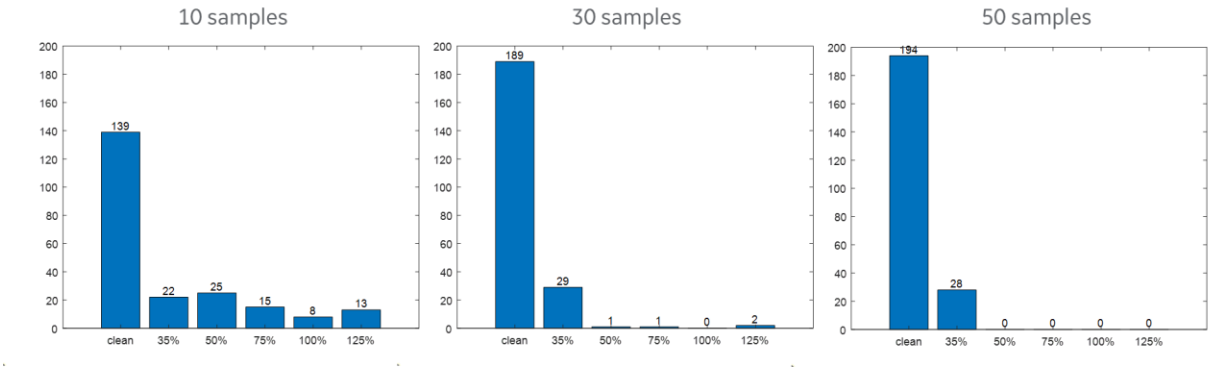



Table 14. Model adaption performance metrics using DOE3a, 10, 30 and 50 samples of 14MW field data.

| # of field sample data | Accurate prediction | Inaccurate prediction |
|---|---|---|
| 10 | 83.8% | 16.2% |
| 30 | 98.6% | 1.4% |
| 50 | 100% | 0 |

## 3.3 Discussions of the observations and findings from Task 3

In this task, our key focus was acquiring field data, calibrating the HIE model built from simulation data, and validating its performance so we can have confidence in implementing the model on real-world turbines. By analyzing GE's Haliade X prototype data collected in Q1 of 2020 and January of 2022, we realized there were two different real-world scenarios that need different calibration strategies and methods. One scenario (Scenario 1) is to deal with the real-world environmental conditions different from the simulation. The second scenario (Scenario 2) is to deal with the individual turbine operation behaviors.

We successfully defined two calibration strategies and methods to tackle each of the scenario respectively. By applying the field data as aforementioned to the model calibration and validation, we proved that:

1) the "calibration via re-simulation" method improved the model performance from 79% to 89% for Scenario 1.

2) the "calibration via model updating" method improved the model performance from 6% to 100% for Scenario 2.

Our findings further validate that our model calibration methods yield satisfactory model performance with available input variables in the real-world. Conclusions from this milestone are crucial in demonstrating how to leverage physics-based simulation data and sparse field data to generate models that can estimate the blade health condition and predict future states, and ultimately enable condition-based maintenance strategy for offshore wind.

Over the course of this project, the Haliade X blade did not appear any leading-edge erosion so the data available for HIE model calibration was from a clean blade. In the future, data from inspecting Haliade X blade will indicate leading edge erosion initiation and development. Future research will include the non-clean blade data in the calibration process and improve the method as needed

# 4 Task 4: Building RUL prediction mode

## 4.1 A suitable roughness growth model identified

In order to plan and schedule the condition-based blade maintenance, knowing only the current blade roughness is not sufficient. The CBM strategy requires predicting the future roughness growth. Our key focus in this effort is to 1) study the prior art of blade leading edge erosion prognostics modeling; 2) identify gaps and challenges in applying the existing methods to the real-world wind turbine and farm operations; 3) define our prognostics approach that can feasibly forecast the future states of blade roughness by tackling the challenges.

### 4.1.1 State of the art prognostics methods

In the past deliverables we demonstrated that it is possible to build a Health Index Estimator (HIE) model for roughness quantification by utilizing standard SCADA data channels. Given the model is reliably able to discern between different roughness states, we expect to detect the onset of blade roughness at an early stage. While it is unlikely that any pertinent action would be needed at that detection point, the interest would be in periodically querying the system, estimating new roughness state, and tracking the rate of deterioration. Once, a deterioration trend is established with certain confidence, such a trend can be further utilized to forecast blade roughness into the future. This forecast specifically determines when the roughness level is likely to cross a predetermined threshold beyond which it is not advisable to operate the turbine without a maintenance action. In the condition monitoring literature this condition-based forecast is referred to as *Prognostics* [5].

#### *4.1.1.1 Prognostics Methods*

In the condition monitoring literature, there are a number of modeling methods that can be largely classified into four broad categories [6]. Depending on availability of information from the asset as well as prognostics resolution required for optimizing the post-prognostic action, the following methods can be considered.

1. ***Reliability Based Methods*** – these are population based statistical summarization of failure rates observed in the field or through manufacturer test programs. Failure data from the past in existing fleets are fit to parametric distributions such as Weibull, which are used to predict failure rates in the assets of interest. Predictions are typically expressed as Mean-time-to-failure (MTTF) that can be used to plan scheduled maintenance programs. These models do not distinguish between how gently or aggressively an asset was used and only provides statistically guaranteed performance over a fleet of assets.

2. ***Proportional Hazard Models*** – These models carefully analyze loading history of the asset to determine life consumption till date, and then subtract estimated consumed life from reliability-based prediction of total life. These models perform better than reliability-based methods alone, however, do not consider knowledge of asset health from inspection or remote monitoring methods.

3. ***Condition-Based Methods*** – These methods rely on remote condition monitoring to assess current health state of the asset, and then use a degradation propagation model along with expected future usage profile to forecast asset health into the future. So far, these methods have been shown to be most versatile and accurate, however the performance depends on quality of monitoring data and a knowledge of fault/degradation propagation mechanism. A spectrum of physics-based, empirical, or data-driven propagation models have been shown to be useful in this context. As an added advantage these methods also provide best quantitative evaluation of prediction uncertainties that must be factored in risk calculations for downstream maintenance planning and scheduling.

4. ***Operational History Data-based Methods*** – in scenarios where a large amount of operational history data from assets are available, recent developments in machine learning and deep learning methods have shown reasonably comparable performance to condition-based methods. The key advantage is that an explicit fault propagation model per failure/degradation mode is not required. This can be extremely powerful, if sufficiently rich operational history data can be made available and in large amounts.

In this work we intend to develop a condition-based method for prognostics. Our HIE model already accomplishes the first key step in prognostics, i.e., current state estimation. Through sequential estimations of current roughness states, and a model to fit a trend on those states will allow us to forecast growth of roughness. Most importantly, these estimates will be recomputed and revised at a predetermined frequency to track any acceleration or deceleration in roughness growth due to environmental and operational loading conditions on the wind turbine assets.

#### *4.1.1.2 Technical Challenges in Prognostics*

Prognostics or accurate estimation of remaining useful life (RUL) is a challenging problem, sometimes due to the type of information available or the lack of it. It is important to consider these challenges while developing a solution, and accordingly interpret the outcomes in downstream decision making appropriately weighing associated risks. Some of the key challenges are enumerated here.

### Challenge 1: Availability and quality of data

Accuracy of predicted quantity heavily depends on three key factors among many others –

(1) *Quality of measurement data* used for current state (roughness severity) estimation - Quality of measurement data is a function of observability through available sensing and sensing resolution. These are relatively controllable if an asset is designed with appropriate sensors or if these sensors can be added once their need is established.

(2) *Our knowledge (model) of damage state evolution under expected operational and environmental loading* - For offshore wind application, our project has established (through simulation data) that commonly measured SCADA parameters can be effective in estimating roughness severity. However, very limited knowledge exists in literature about roughness growth model over time and almost no operational data exists to our knowledge. Some of these models have been discussed in prior section, which will be leveraged in our simulation of roughness growth curves. Furthermore, some inspection-based efforts within GE intend to document leading edge roughness growth for onshore wind turbines, which are under our consideration to glean this knowledge to the extent possible within time and budget scope of our project.

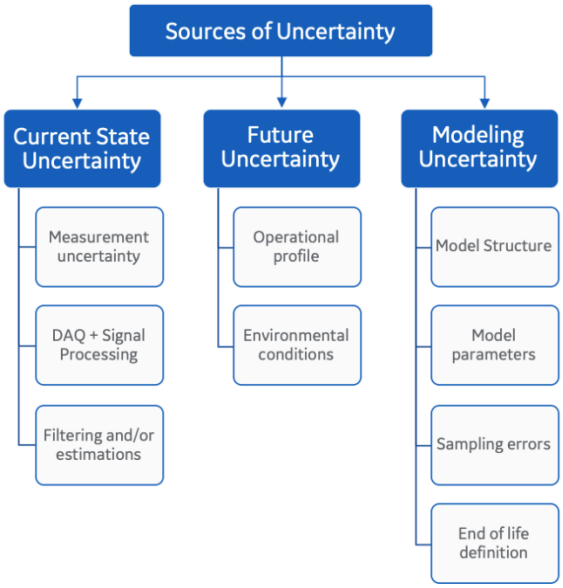
(3) *Accuracy of the knowledge about expected future loading on the system* – Given a damage state, how fast or slow it will evolve largely depends on future stress vectors such as environmental and operational loading, which can at best be forecasted with associated uncertainties. In addition, knowing the actual failure mechanism is critical in understanding how quickly the material or component will fail. Examples of this include material erosion, defect driven cracks and delamination. Prognostic methods tend to forward propagate a range of scenarios to produce distributions of expected degradation, however the accuracy of any forecast heavily depends on accuracy of weather forecast as well as accuracy of operational loading forecast.

As mentioned earlier, while quality of measurement is sometimes controllable, the other two factors (i.e. 2 and 3) are often learned through known domain physics (i.e. a physics model of a crack growth over time) or from historical data, if made available in sufficient quantities (e.g. several run-to-failure trajectories, or weather patterns at a site). Both these factors have been challenging in the real-world offshore wind as well and we intend to utilize as much information available as possible and make some assumptions where needed.

### Challenge 2: Uncertainties

There are several sources of uncertainties that play an important role in prognostic estimations (see Figure 20). To briefly enumerate, these sources of uncertainty include – measurement data (present state) uncertainty, future loading/usage uncertainty, modeling uncertainty (both state estimation and forecasting methods), all of which tend to cumulatively stack up as long-term forecasts are generated [5]. Hence, a forecast far out in future tends to have higher uncertainty than a relatively short-term prediction, even if the sources of uncertainties remain the same.

Figure 20. Various sources of uncertainties affecting accuracy of prognostics (adapted from [5]).




## 4.1.2 Challenges in in building blade roughness growth model for offshore wind

In addition to the technical challenges in building prognostics models in general, this section will summarize the unique challenges in building blade roughness growth model for offshore wind by translating the aforementioned technical challenges to this project.

### *4.1.2.1 Scarce real-world data*

To build a blade roughness (or LEE) growth model for offshore wind, either physics-based, empirical, or data-driven propagation models can be applied as a condition-based prognostics modeling method. In this context of offshore wind, limited physical knowledge and lacking real-world data are the challenges.

#### *4.1.2.2 Definition of remaining useful life*

Another challenge is about how to define the remaining useful life for blade roughness. Different from structural degradations/failures impacting the blade health, leading edge erosion changes blade surface roughness, which thereafter degrades blade's aero performance. However, leading edge erosion may not necessarily cause imminent blade structural failure that is conventionally defined as the so-called "end-of-life". As described in our proposal, the ultimate industrial value of the leading-edge erosion growth model is to enable CBM to reduce the OPEX while recovering AEP loss due to aero performance degradation from LEE. Therefore, instead of using "blade structural failure", we choose a maximum allowable percentage of AEP loss before repair as the definition of "end-of-life" and define "remaining useful life" as the remaining time to the end of life in our roughness growth model. This maximum allowable percentage of AEP loss will depend on the farm owner's overall O&M strategy and other constrains. By no means we could suggest a universally representative "remaining useful life" for blade roughness. As one example and for the purpose of technology development, we tend to define 1.5% of AEP loss due to leading edge erosion as the "end-of-life" for our roughness growth model. Since the offshore wind industry is still in the exploratory phase where new technologies are being developed and the farm-level CBM strategy is being built, outcomes from this project will be a critical part of that strategy. Our goal of this milestone is to develop a POC blade roughness growth model under the assumption of 1.5% AEP loss being considered as the "end-of-life".

#### *4.1.2.3 Compounding effect of lightening on blade roughness change*

The last challenge is to decouple the blade roughness changes due to a catastrophic damage event such as a lightning strike from the roughness growth due to erosion damage from rain and hail. Lightning strikes on offshore wind blades is a serious concern due to its potential damages on blade health and performance. To ensure our roughness growth model can accurately forecast the future states, we need an approach to monitor and detect the impact of lightening events and update the roughness growth model if necessary.

#### *4.1.2.4 Correlating blade roughness to observable LEE in real world*

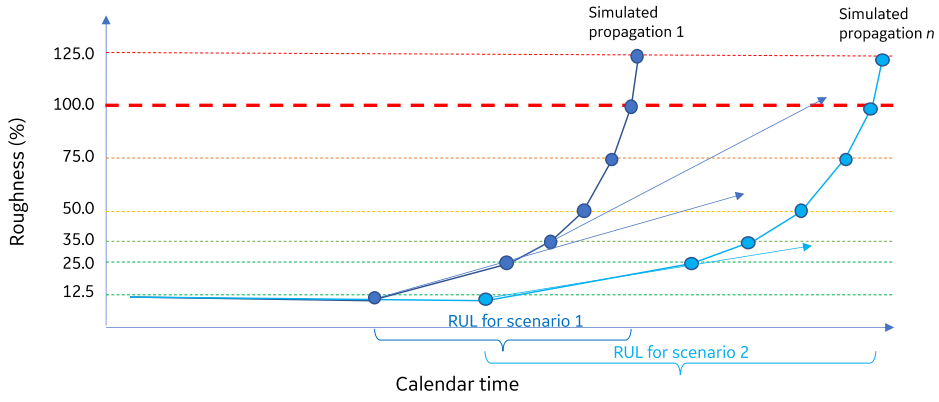
As described in Task 1, to build a blade HIE model without a large amount of field data, we leveraged a high-fidelity Bladed physical model to generate simulation data for Haliade X. Since the blade leading edge health state (ie. erosion damage or roughness) per say is not an input parameter in Bladed, we instead represented this by modification of aerodynamic lift and drag coefficients at different blade sections as inputs in Bladed which is a result of erosion induced roughness changes along the leading blade edge. However, blade roughness changes are not a directly measurable parameter today in the real-

world O&M practice or lab testing for quantifying the blade LEE growth. Instead, visual inspection to observe the physical material damages such as pit, dent, gouge, and delamination is performed to monitor the erosion growth. To validate a roughness growth model with real-world leading edge erosion data, one primary challenge is to map the observed blade LEE quantifications to roughness changes that we used in our models.

### 4.1.3 Blade LEE growth model and validation approaches

Previous milestones demonstrated that ML-trained HIE model can effectively detect and ascertain the degree of roughness from simulation data. To demonstrate this capability, simulations were run with aerodynamic coefficients that are expected to represent various levels of roughness ranging from clean (0-12.5%) to very rough (>125%), relative to a 100% roughness that defined turbine trip setpoint through wind tunnel tests. One key piece of information that is missing is the time evolution of roughness, i.e. for instance, how long it could take for roughness to grow from ~12.5% to 25% as compared to 25% to 50%.

Figure 21. Illustrative roughness growth trends using simulation




As described in reference [7], previous works established a linear relationship between environmental stresses (measured in cycles) and mass loss. Mapping of cycles to calendar time is required to integrate prognostic information into maintenance processes. Given time-scales (months-years) for such roughness to grow, it may be a reasonable assumption to model linear growth, however, previous studies on a number of physical systems have shown that typically fault growth tends to follow exponential growth patterns [8]. In absence of operational data from actual offshore farms and any experience with leading edge erosion patterns, we will develop a generalized prediction model with capability to progressively learn any arbitrary monotonic growth pattern illustrated in Figure 21. We will use state-of-the-art condition-based methodology that follows a two-step recursive estimation approach mentioned in Section 4.

## 4.2 Completion of the initial prediction model and data preparation; Completion of RUL model building and validation

### 4.2.1 Data generation for RUL model

In the past deliverables we demonstrated that it is possible to build a Health Index Estimator (HIE) model for roughness quantification by utilizing standard SCADA data channels. Given the model is reliably able to discern between different roughness states, we expect to detect the onset of blade roughness at an early stage. From the onset, an RUL model is needed to forecast blade roughness into the future. This forecast specifically determines when the roughness level is likely to cross a predetermined threshold beyond which it is not advisable to operate the turbine without a maintenance action. In the condition monitoring literature this condition-based forecast is referred to as *Prognostics* [5]. In this report, RUL model and prognostics model are exchangeable terminology.

#### *4.2.1.1 Relationship between roughness and AEP loss*

Throughout this program, we use blade roughness to quantify the blade health status and define the end-of-life as the time when the roughness level exceeds a predetermined threshold. However, as a common industrial practice, rain erosion lab tests measure AEP loss to study the change of blade health over time. Therefore, we establish the relationship between blade roughness and AEP loss so that the change of blade roughness over time can be assessed.

Our previous Bladed simulation has generated power curve for several specified roughness levels and the corresponding AEP can be calculated. The AEP for the initial roughness level of 12.5% is treated as a reference AEP and accordingly the AEP loss is 0. Higher AEP loss can be caused at larger roughness level, which is shown in Figure 22(a). To enable evaluation of the roughness level for any specific AEP loss, we develop a nonlinear model to characterize the relationship between them.

First, we transform roughness values using the Box-Cox method so that the transformed roughness values are more normally distributed and potentially more linearly correlated with AEP loss.

$$y^* = f(y) = \begin{cases} \frac{y^l - 1}{l}, & if\ l \neq 0, \\ \log(y), & if\ l = 0, \end{cases}$$ where $y$ is roughness.

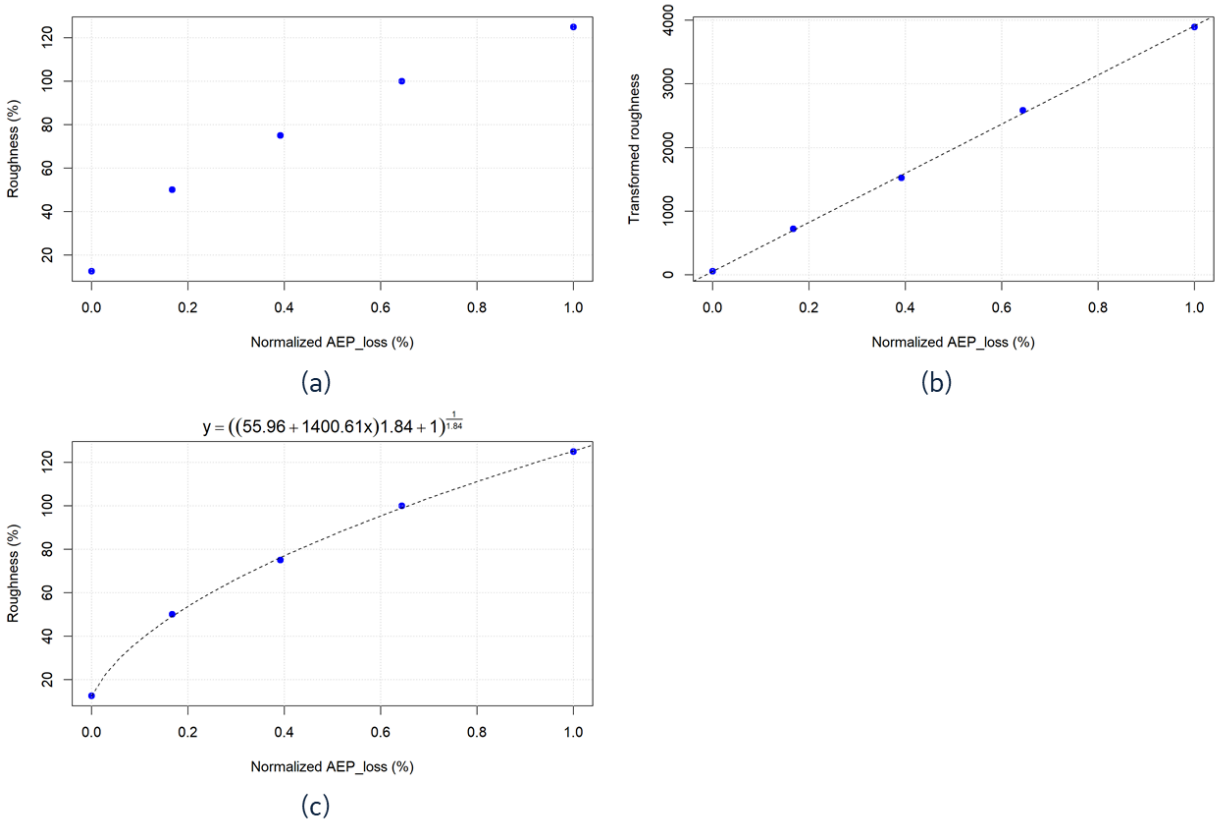
The maximum likelihood estimates of $l$ is 1.84 and the resulting transformation is $y^* = (y^{1.84} - 1)/1.84$. Second, for the transformed roughness vs. AEP loss shown in Figure 22(b), we fit a linear regression

model with the constraint that the fitted linear line must go through the initial point. The fitted linear model using the maximum likelihood method is $y^* = 55.96 + 1400.61x$. Combining results from the above Box-Cox transformation and linear regression model, the relationship between roughness and AEP loss is characterized by the following equation, which is also shown in Figure 22(c):

$$y = [1.84(55.96 + 1400.61x) + 1]^{1/1.84},$$

where $y$ is roughness and $x$ is AEP loss.

Figure 22. Relationship between roughness and AEP loss

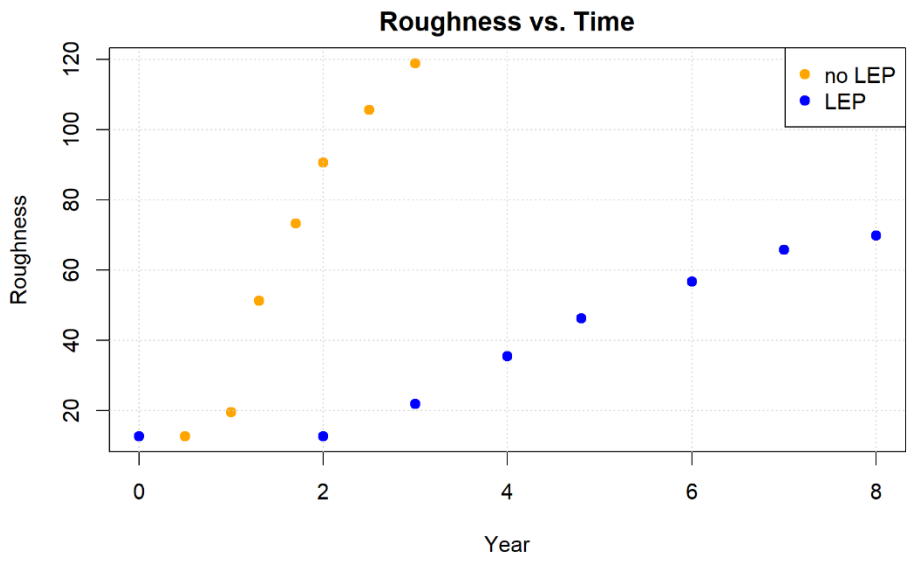



### *4.2.1.2 Lab-based rain erosion test data*

For research purpose, a GE team has carried out a rain erosion lab test to measure how AEP changes over time due to blade erosion. In this lab test, the blade with and without the leading edge protection (LEP)

system were both considered, and a constant rain intensity was used. LEP is a piece of clamp-on compoennt that protects leading edge from erroded. Figure 23 shows the roughness change over time based on the lab test using the roughness vs. AEP loss relationship deveoped in the previous section. An incubation time period is present for both blades with and without LEP, which is two years and half year, respectively. After the incubation time period, the roughness value for the blade without LEP increases much faster than that with LEP, indicating that the LEP system can protect blade efficiently and extend blade life.

Figure 23. Lab test data




To facilitate the data generation discussed in the next section, we develop a baseline RUL model to characterize the increasing trend of roughness over time given the constant rain intensity that was used in the lab test. Specifically, we consider a constrained nonlinear model as follows:

$$y = y_0 + a(x - x_0)^b + \varepsilon,$$

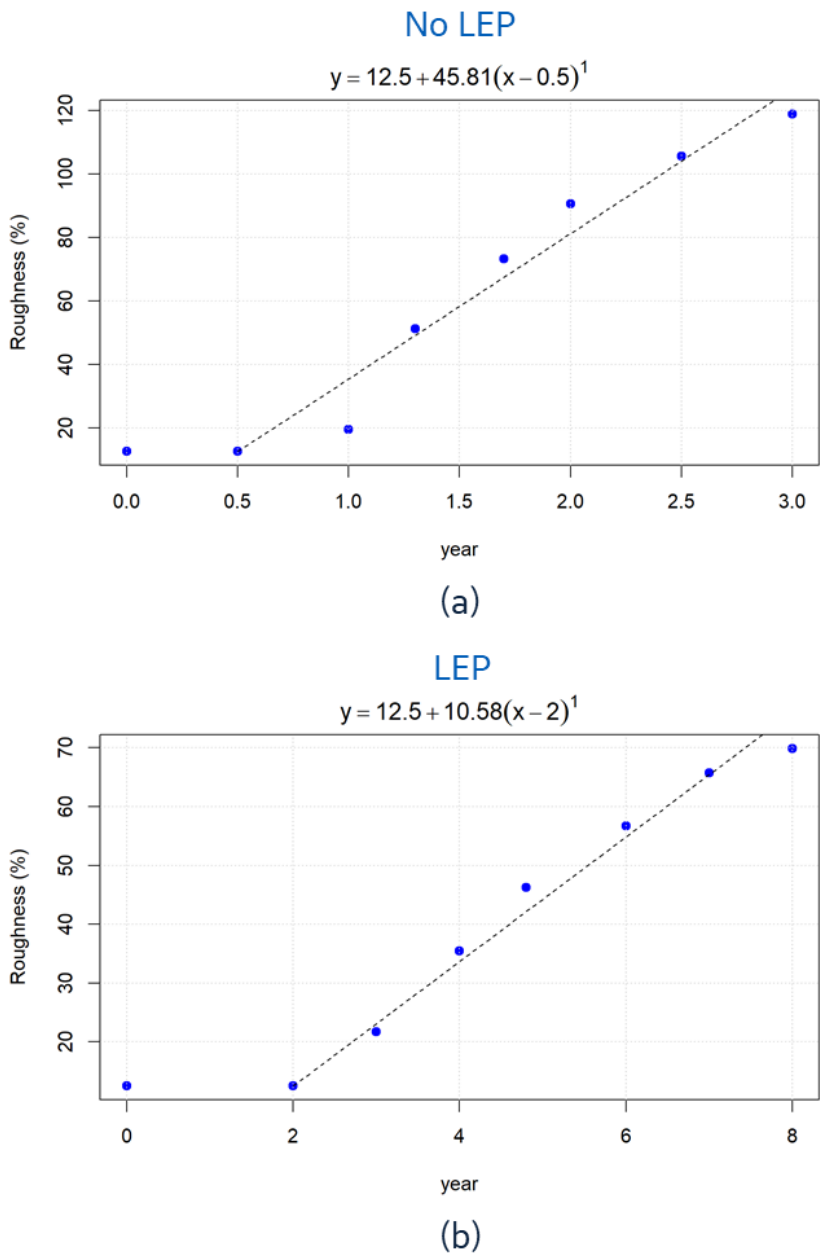
where $x$ is time (year), $y$ is roughness (%), $x_0$ is the incubation time, $y_0 = 12.5$, $a > 0$, $b \geq 1$, and $\varepsilon$ is the error term. This nonlinear model implies that the fitted curve must go through the initial point $(x_0, y_0)$, it is monotonically increasing, and its increasing rate can potentially increase over time. Note that the linear trend is a special case of the above nonlinear model when $b = 1$ meaning the increasing rate is a constant. As shown in Figure 24, the fitted baseline RUL models for the blade with and without LEP are:

$y = 12.5 + 10.58(x - 2)$ for the blade with LEP,

$y = 12.5 + 45.81(x - 0.5)$ for the blade without LEP.

These two fitted models indicate that a linear trend represents how roughness changes over time after the incubation time period based on the limited lab test data.

Figure 24. Baseline RUL model



### *4.2.1.3 Data generation originated from lab-based rain erosion test*

The change of blade roughness over time can be impacted by multiple factors such as precipitation, wind speed, and turbine operation. Here, we combine the baseline RUL model and the precipitation data at Rotterdam, the Netherlands to simulate data that represents how blade roughness may change over time in the real world.

Figure 25(a) shows the monthly precipitation data at Rotterdam, the Netherlands since 2010, and Figure 25(b) presents its distribution normalized by the constant rain intensity used in the lab test. Our data generation process is summarized in Table 15.

Figure 25. Precipitation at Rotterdam

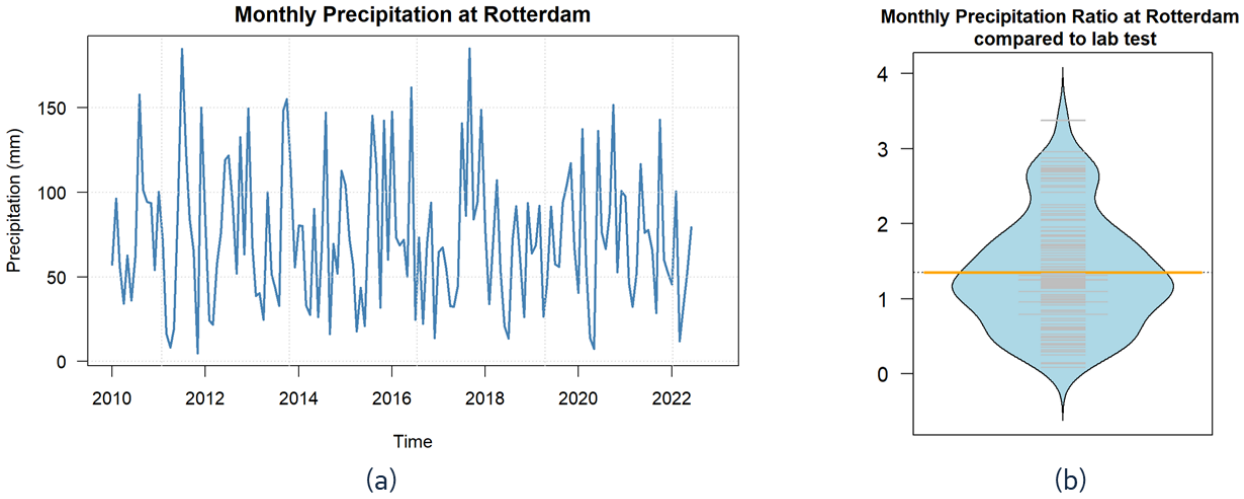
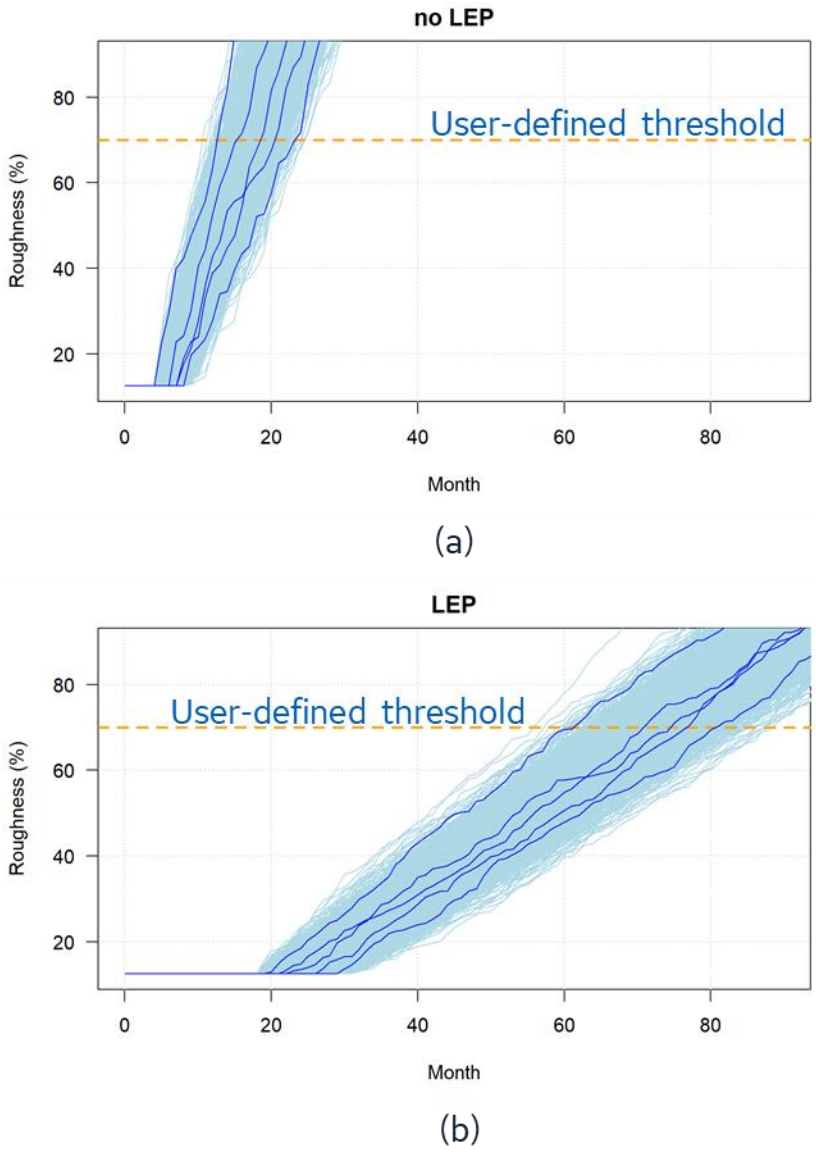



Table 15. Data generation process

| |
|---|
| Step 1: Start with time 0 for which the initial roughness level is $y_0 = 12.5\%$. |
| Step 2: Randomly generate an incubation time $x_0 \sim Uniform(x_{0,min}, x_{0,max})$. Specify all the roughness values until $x_0$ the same as the initial roughness level $y_0 = 12.5\%$. |
| Step 3: For each time point after $x_0$, say $x_i$, calculate the increasing rate $a_i$ starting from its previous time point $x_{i-1}$, which is equal to the increasing rate $a$ in the baseline RUL model multiplied by a random sample of the precipitation ratio. |
| Step 4: Calculate the roughness level at time $x_i$ which is $y_i = y_{i-1} + a_i(x_i - x_{i-1})$. |
| Step 5: Repeat Step 3-4 and sequentially generate data points $(x_i, y_i)$ until a pre-determined maximum time point. |

Here, the data points $(x_i, y_i)$ are generated monthly. For the blade with LEP, the baseline increasing rate $a = 10.58/12$ and $x_0 \sim Uniform(18, 30)$. For the blade without LEP, the baseline increasing rate $a = 45.81/12$ and $x_0 \sim Uniform(4, 8)$. We repeat the data generation process 1,000 times, and all the generated 1,000 curves are shown in Figure 26. Given a user-defined threshold (i.e, 70% roughness), it can be observed that the blade without LEP will reach to the end of life much faster than that with LEP, which reflects the effectiveness of the LEP system.

Figure 26. Data generation summary



### 4.2.2 RUL model development and validation

The generated data discussed in Section 4.2.1.3 represents a simulated ground truth of how blade roughness level increases over time and when it will reach to a user-defined threshold. At any time, only partial data of roughness vs. time is available which will be used to build an RUL model to predict remaining useful life before the threshold is reached. In this section, we discuss the RUL model development and perform model validation.

#### *4.2.2.1 RUL model building*

An RUL model can be developed at any time when it passes the incubation period and the available roughness vs. time data is sufficient for model building. An example is shown in Figure 27(a), where the roughness vs. time data is available for the first 11 months and the incubation period is 6 months. Based on the simulated ground truth, the true end of life in this case is the 22$^{nd}$ month when roughness exceeds a given threshold of 70%. Hence, the true RUL at the modeling time of the 11$^{th}$ month is 22-11=11 months.

To build an RUL model, we consider two functional forms to characterize the increasing trend of roughness over time. The first one assumes linear trend as follows:

$$y = y_0 + a(x - x_0) + \varepsilon,$$

where $x$ is time (month), $y$ is roughness (%), $x_0$ is the incubation time, $y_0 = 12.5$, $a > 0$, and $\varepsilon$ is the error term. The model implies that the fitted linear line will go through the initial point $(x_0, y_0)$ and it is monotonically increasing.

The second one allows non-linear trend as follows:

$$y = y_0 + a(x - x_0)^b + \varepsilon,$$

where $x$ is time (month), $y$ is roughness (%), $x_0$ is the incubation time, $y_0 = 12.5$, $a > 0$, $b \geq 1$, and $\varepsilon$ is the error term. This nonlinear model implies that the fitted curve will go through the initial point $(x_0, y_0)$, it is monotonically increasing, and its increasing rate can potentially increase over time. Note that the linear trend is a special case of the above nonlinear model when $b = 1$ meaning the increasing rate is a constant.

Figure 27. An RUL model building example

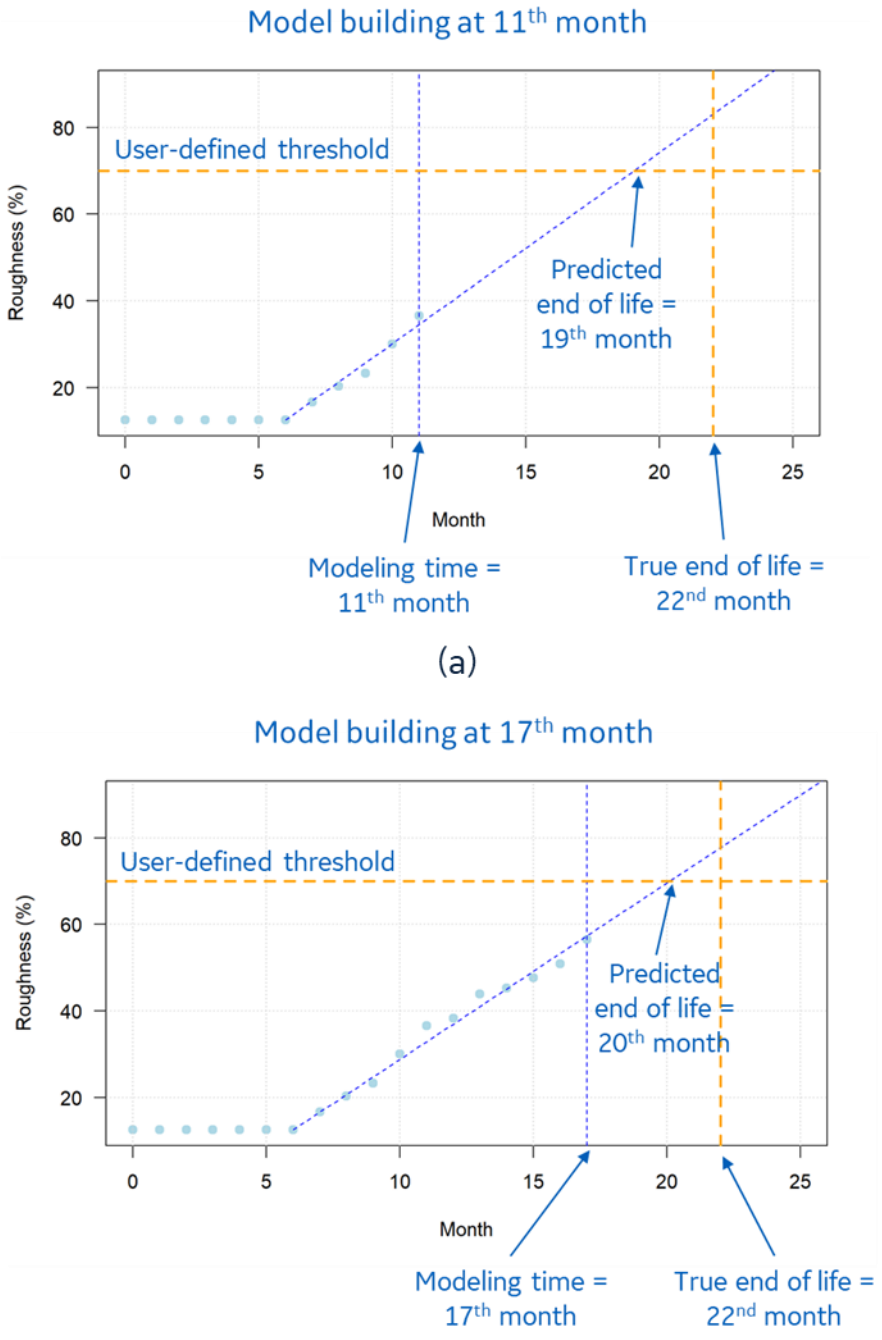



Figure 27(b) illustrates that a linear RUL model is built at the 11th month for a blade without LEP. Applying this model to the future time points, we can measure that the predicted end of life is the 19th month when the predicted roughness exceeds the threshold of 70% and therefore the predicted RUL is 19-11=8 months. Compared with the true RUL of 11 months, this model underestimates the RUL by 11-8=3

months. The same RUL model building process can be performed repeatedly as time moves on. Figure 27(b) shows that a linear RUL model is re-built at the 17$^{th}$ month as a few more data points become available. Correspondingly, the predicted end of life based on this updated RUL model is the 20$^{th}$ month, and the predicted RUL is 20-17=3 months. The true RUL at this new modeling time is 22-17=5 months, therefore the re-built model underestimates the RUL by 5-3=2 months. Notice that the updated RUL model at the 17$^{th}$ month performs better than that at the 11$^{th}$ month as the updated predicted end of life is closer to the true end of life and the RUL prediction error is smaller, which indicates that the RUL model built upon more recent data points can better capture the increasing trend of roughness over time and provide more accurate RUL prediction.

#### *4.2.2.2 RUL model validation*

To perform the RUL model validation and evaluate model performance, we apply the RUL model building process to each of the 1000 generated data curves. For each data curve, the RUL model is updated at every time point starting with an after-incubation time when a sufficient set of data points are available for model building, and the updated RUL model is used to predict the end of life and RUL. Also, both linear and nonlinear RUL models are considered.

The model validation results for the blades with and without LEP are summarized in Figure 28 and Figure 29, respectively. Each summary plot shows the change of RUL prediction error against the true RUL at the modeling time (the difference between the modeling time and the true end of life). For each true RUL (or the number of months prior to the end of life), a boxplot is used to summarize the distribution of RUL prediction error across the 1000 curves. Each boxplot includes a box displaying the median, the 1$^{st}$ quartile and the 3$^{rd}$ quartile, a lower whisker which is the 1$^{st}$ quartile minus 1.5 times the interquartile range, and an upper whisker which is the 3$^{rd}$ quartile plus 1.5 times the interquartile range.

It can be observed that the variation of the distribution of RUL prediction error is smaller as the modeling time gets closer to the true end of life, which indicates that the updated RUL model can leverage more recent data to better capture the roughness change over time and predict RUL more accurately. Furthermore, the nonlinear model allows the increasing rate of roughness to potentially increase with time. Hence, when the trend based on partial data at a modeling time is significantly different from the trend that would be represented by the entire data, the predicted end of life from the nonlinear model is likely to be earlier than the true end of life and the nonlinear model tends to underpredict RUL. This can happen more likely at a very early modeling time leading to a larger RUL prediction error for the nonlinear model compared to the linear model as shown in Figure 28.

Figure 28. Model validation for LEP

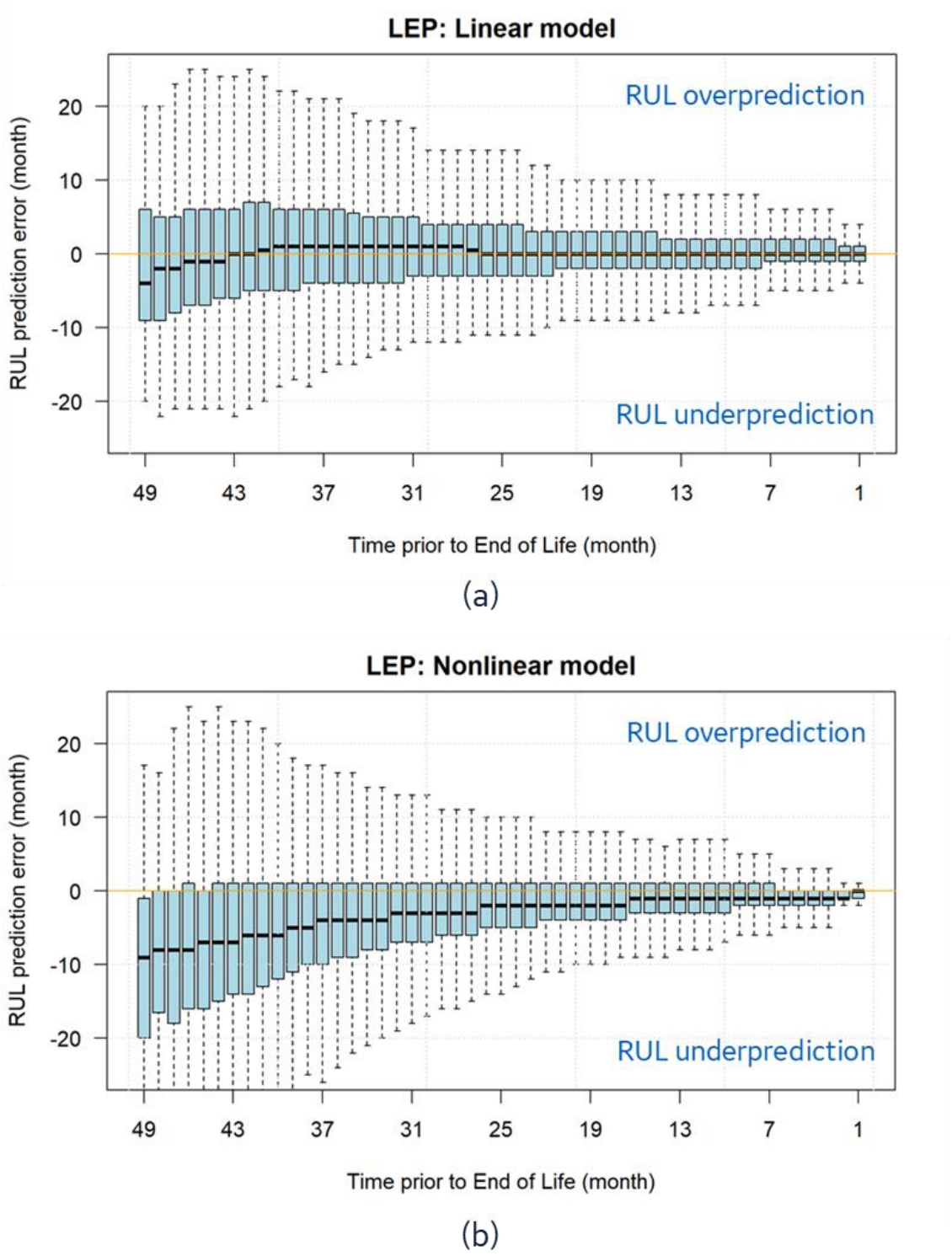
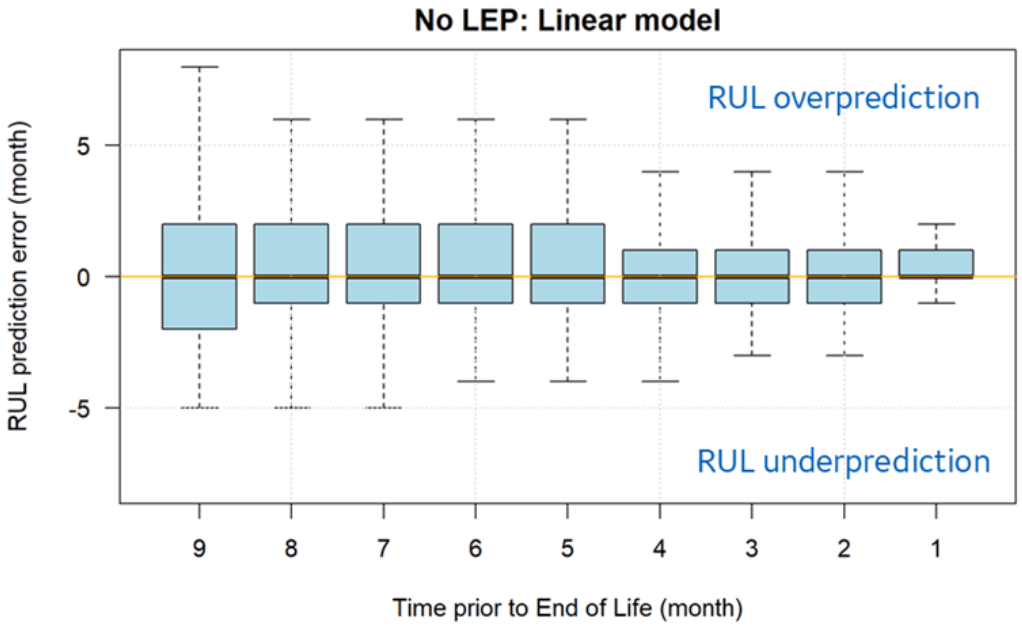
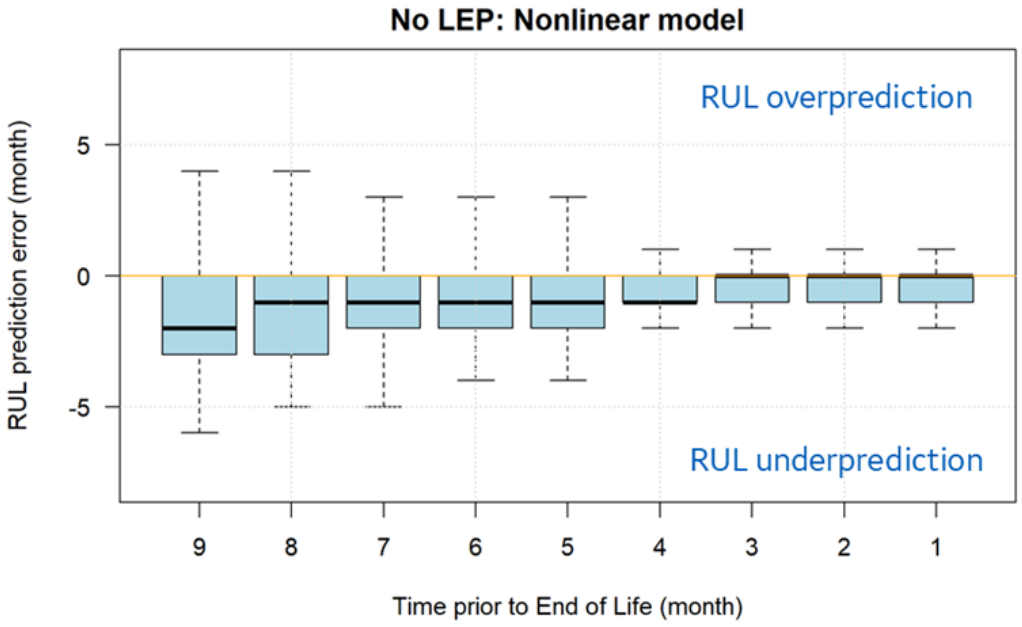


Figure 29. Model validation for no-LEP




(a)




(b)

## 4.3 Discussions of the observations and findings from Task 4

In this Task, we firstly investigated the state-of-the-art prognostics modeling methods and defined the challenges in applying those methods for offshore wind. We then studied a representative rain erosion lab test data base on which a baseline RUL model was developed to characterize the increasing trend of roughness over time given a constant rain intensity. Combining the baseline RUL model with the precipitation data in Rotterdam where GE's Haliade X prototype was installed, we simulated a set of data to represent how blade roughness may change over time in the real world. For the generated data, we developed a roughness RUL model building process and performed model validation to demonstrate the effectiveness of our proposed modeling method.

More studies can be performed in the future. For example, when other factors are considered in the lab test such as wind speed and turbine operational conditions, we can incorporate those factors into our data generation and RUL modeling process to further understand whether and how blade roughness change over time would be impacted. To map the real-world blade leading edge erosion conditions to roughness level changes, significant research and testing efforts are required in the future.

# 5 Project summary

**Contributions and results:**

In this project we have built and proved a framework for developing HIE and RUL models to realize CBM when historical data and failure data are limited. This framework includes 1) using physical model as simulator to generate DOE data to complement the scarce real-world data; 2) using most advanced machine learning methods to build health index models; 3) calibrating the health index models with available field data so the models can be deployable; 4) propagating the health index to the future with a RUL model to enable CBM. As proof-of-concept, our use case is blade leading edge erosion. We proved this framework technology based upon GE's Haliade X offshore turbine blade. We believe this framework is generally applicable for other failure modes as well as different turbine components.

**Remaining challenges and our recommendations:**

Certainly, more studies can be performed in the future on the RUL modeling area. For example, when other turbine operation factors are considered in the lab test such as wind speed and rotor speed, we can incorporate those factors into our data generation and RUL modeling process to further understand whether and how these factors would impact the blade roughness changes over time. Since GE's offshore turbines and fleets are still young, there are no noticeable leading-edge erosions or roughness changes yet, which makes the large-scale field validation for HIE and RUL models impossible at this time. When more field data becomes available, we will further study and enhance our calibration methods.

**Future efforts:**

As one of the future effort, GE Research team is working with GE Vernova's offshore business to deploy this HIE model to the Haliade X prototype for further validation, and plan for large scale validation when other large-scale projects come to live. In the meanwhile, we intend to apply this framework technology to address the HIE and RUL for other defects critical to offshore wind turbine maintenance, such as blade crack and blade delamination. The O&M cost reduction using CBM for blade will be realized once a holistic solution is built, validated, and implemented to tackle majority of the failure modes.

# 6 Bibliography

[1] G. Huang, etc., "Extreme learning machine: Theory and applications," *Neurocomputing,* pp. 489-501, 2006.

[2] G. Huang, etc., "Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks," *Proceedings of International Joint Conference on Neural Networks (IJCNN2004),* 2004.

[3] G. Huang, "An Insight into Extreme Learning Machines : Random Neurons , Random Features and Kernels," New York, Springer New York LLC, 2014, pp. 376-390.

[4] GE Research, "154630_GE_deliverable_D0A.1_M3," 2021.

[5] K. Goebel, M. J. Daigle, A. Saxena, I. Roychoudhury, S. Sankararaman and J. R. Celaya., Prognostics: The science of making predictions., Createspace Independent Publishing Platform, 2017, p. 396.

[6] A. Saxena, S. Sankararaman and K. Goebel, "Performance Evaluation for Fleet-based and Unit-based Prognostic Methods," in *2nd European Conference of the PHM Society*, Nantes France, 2014.

[7] A. Dashtkar and e. al, "Rain erosion-resistant coatings for wind turbine blades: A review," *Polymers and polymer composites,* 2019.

[8] K. Goebel, M. J. Daigle, A. Saxena and e. al, Prognostics: The science of making predictions, Createspace Independent Publishing Platform, 2017.

[9] KNMI, "KNMI data platform," Royal Netherland Meteorologic Institute, [Online]. Available: https://dataplatform.knmi.nl/organization/.

[10] B. Saha, K. Goebel, A. Saxena and J. Celaya.US Patent 8,725,456, 2014.

[11] A. Saxena, J. S. B. Celaya and S. G. K. Saha, "Metrics for Offline Evaluation of Prognostics Performance," in *International Journal of Prognostics and Health Management (IJPHM)*, 2010.

[12] IEC, "64100-1: Wind Energy Systems -- Part I: design requirements," IEC, 2019.

[13] F. Wilcoxon, "Individual Comparisons by Ranking Methods," *Biometrics Bulletin,* vol. 1, no. 6, pp. 80-83, 1945.

[14] H. B. a. D. R. W. Mann, "On a Test of Whether One of Two Random Variables Is Stochastically Larger Than the Other," *The Annals of Mathematical Statistics,* vol. 18, no. 1, pp. 50-60, 1947.

[15] GE Research D2.1, "154630_GE_deliverable_D2.1_M8," Niskayuna, 2021.

[16] G. Springer and a. C. Bax, A Model of Rain Erosion for Homogeneous Materials, AFML-TR-72-106 (Air Force Materials Directorate).

[17] G. Yang, S. Springer and C. I., "Model for the Rain Erosion of Fiber Reinforced Composites," *AIAA,* vol. 13, no. 7, p. George S. Springer and Cheng I. Yang, 1975.

[18] e. a. Arash Dashtkar, "Rain erosion-resistant coatings for wind turbine blades: A review," *Polymers and polymer composites,* 2019.

[19] H.M.Slot, E.R.M.Gelinck, C.Rentrop and E. d. Heide, "Leading edge erosion of coated wind turbine blades: Review of coating life models," *Renewable Energy,* vol. 80, pp. 837-848, 2015.

[20] A.N.Kaore, U.B.Kale, C.S.Yerramalli and H.K.Raval., "Turbine specific fatigue life prediction model for wind turbine blade coatings subjected to rain erosion," *Materials Today Communication,* vol. 31, 2022.

[21] L. Mishnaevsky Jr, "Repair of wind turbine blades: Review of methods and related computational mechanics problems," *Renewable Energy,* 2019.

[22] J. Chen and J. Wang, "Review on rain erosion protection of wind turbine blades". Journal of Coatings Technology and Research," *Journal of Coatings Technology and Research,* vol. 16, no. 1, pp. 15-24, 2019.

[23] B. Amirzadeh, A. Louhghalam, M. Raessi and M. Tootkaboni, "A computational framework for the analysis ysis of rain-induced erosion in wind turbine blades, part i: Stochastic rain texture model and drop impact simulations," *Journal of Wind Engineering and Industrial Aerodynamics,* vol. 163, pp. 33-43, 2017.

[24] GE Research, "154630_GE_deliverable_D1.3_M6," 2021.

[25] G. Research, "154630_GE_deliverable_D1.3_M6," Niskayuna, 2021.

[26] GE Research, "154630_GE_deliverable_D1.3_M6," 2021.

[27] R. GE, "154630_GE_deliverable_D2.2_M9."," 2022.

[28] GE Research, "154630_GE_deliverable_D0A.1_Q2_M6," 2021.