Investigation of Multiple Data Streams for Gearbox Bearing Fault Prediction Through Machine-Learning Models

Lindy Williams, Yi Guo, Shawn Sheng,
National Renewable Energy Laboratory,
and Arch Desai, Spark Cognition

Wind Power Data and Digital Innovation
October 29, 2021

Outline

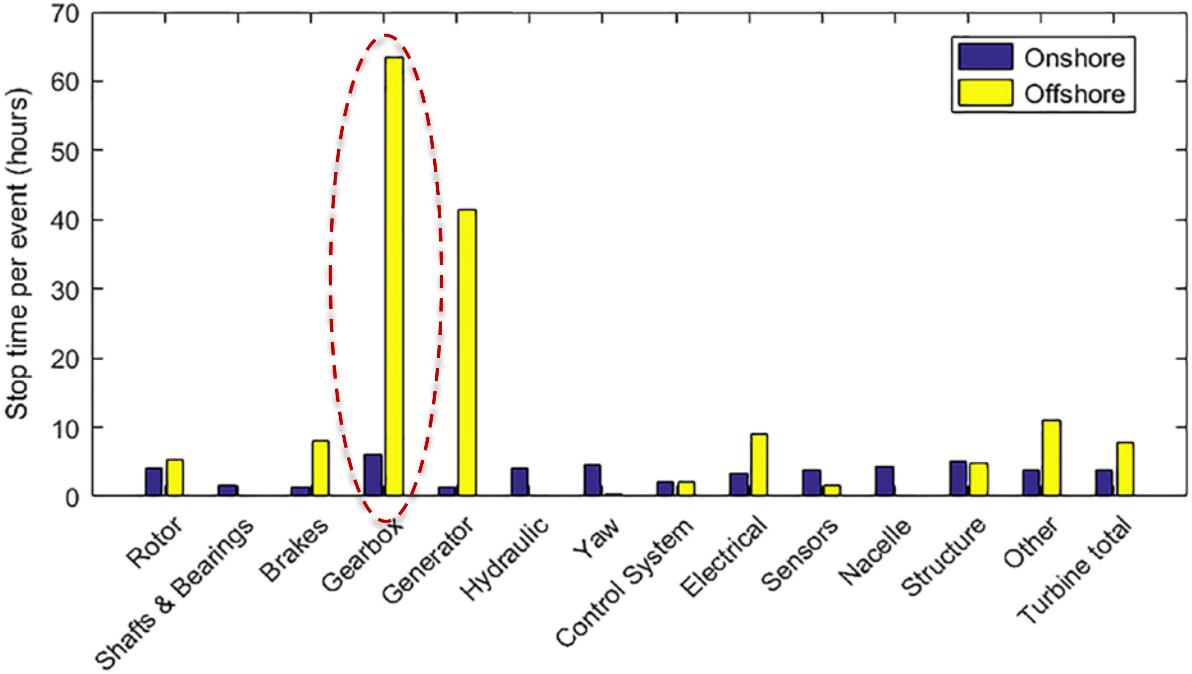
- Background
- Methodology
- Case Studies
  - 1: Supervisory Control and Data Acquisition (SCADA) and Modeled Data
  - 2: SCADA, Modeled Data, and Condition-Monitoring Data
- Summary

U.S. Department of Energy 1.5-MW turbine, NREL 17702

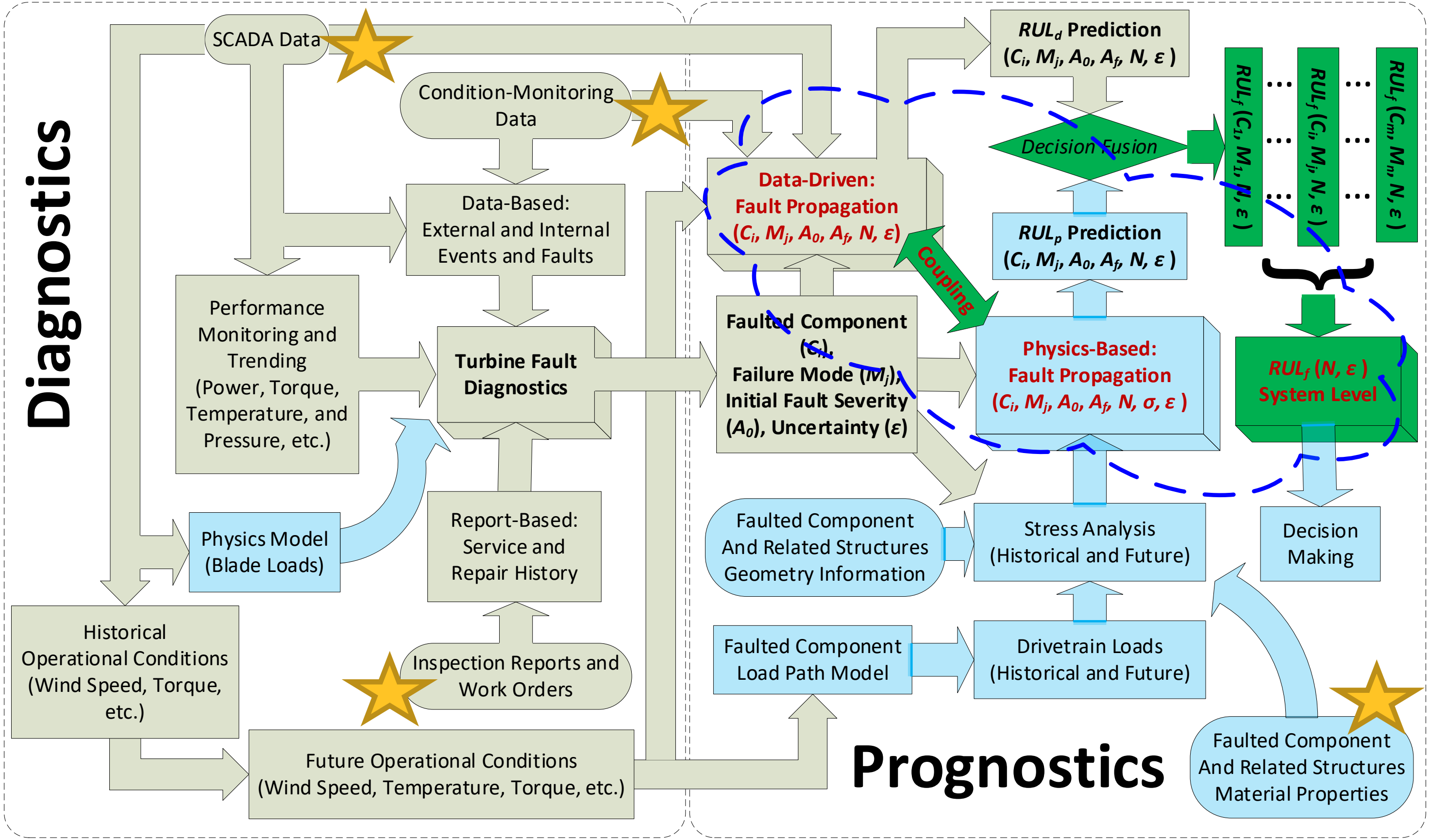
Background

Gearbox Reliability Challenge

- Eighteen publicly available databases were reviewed by colleagues at Durham University.[1]
- Stop time per event, led by gearboxes, challenges the wind industry and results in an increased cost of energy for wind power.

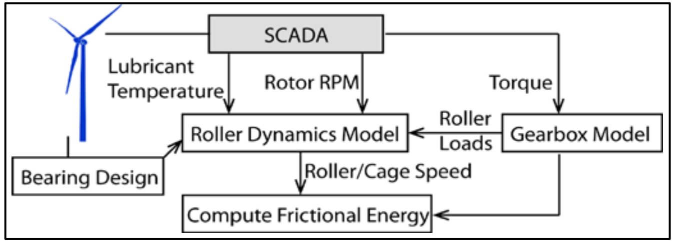


A "Prognostics and Health Management for Wind" Framework[2]

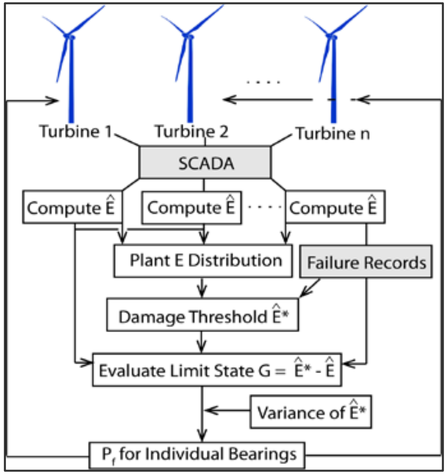


Gearbox Bearing Fault Prediction

- Target top failure mode: bearing axial cracking
- A **reliability assessment and prognosis methodology** through integration of physics domain modeling and data domain inputs[3]
  - Individual bearing probability of failures throughout entire life cycle
- **Objective**: **To investigate multiple data streams**, including SCADA, physics domain modeled data (generated using physics domain models of the reliability assessment prognosis methodology mentioned on previous slide), and condition-monitoring data, **for gearbox bearing fault prediction through machine-learning (ML) models.**[4]



Modeling steps for calculating frictional energy of individual bearings

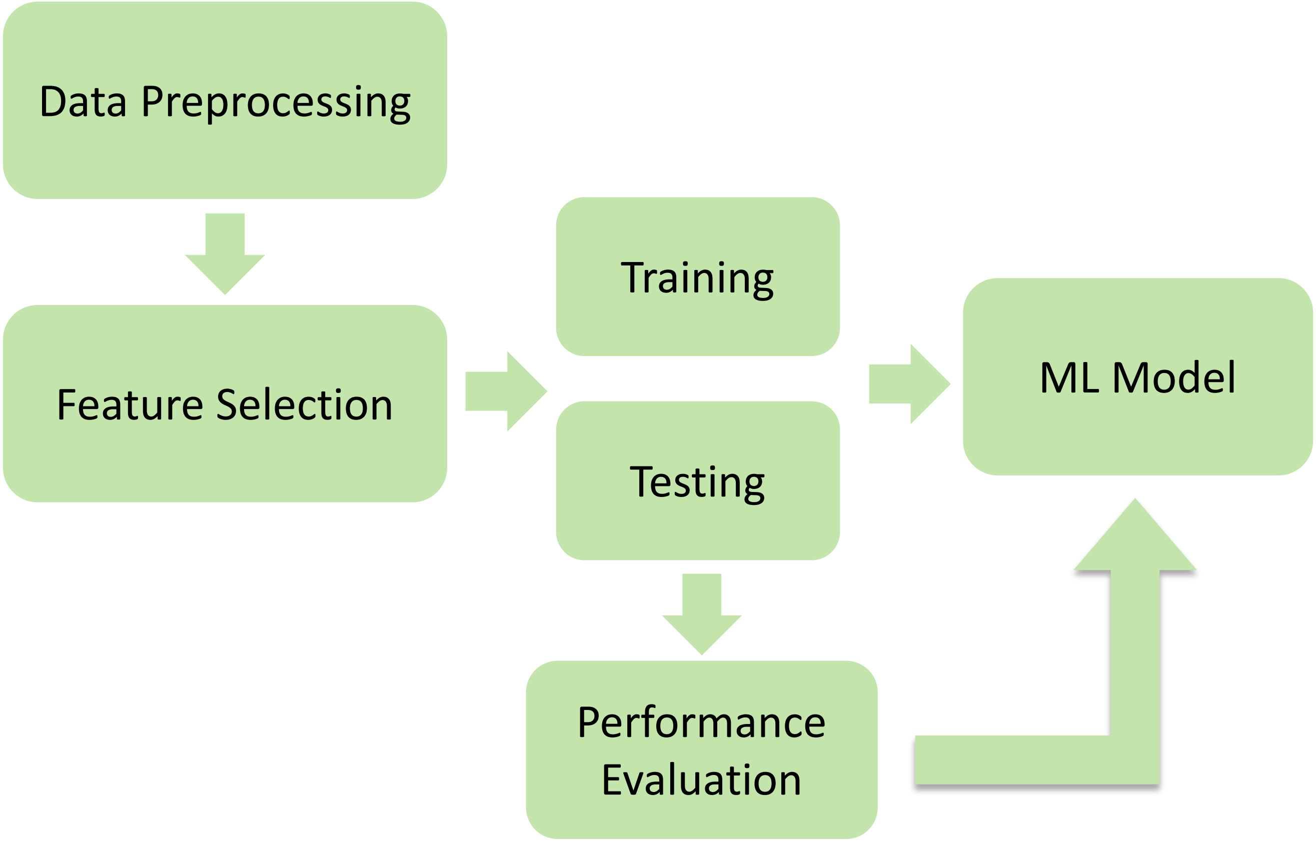


E: frictional energy
$\hat{E}$: nondimensional frictional energy
$\hat{E}^*$ threshold for $\hat{E}$
G: limiting function

Modeling steps for calculating the probability of failure ($\boldsymbol{P_f}$) of individual bearings

Methodology

Machine-Learning Model Development



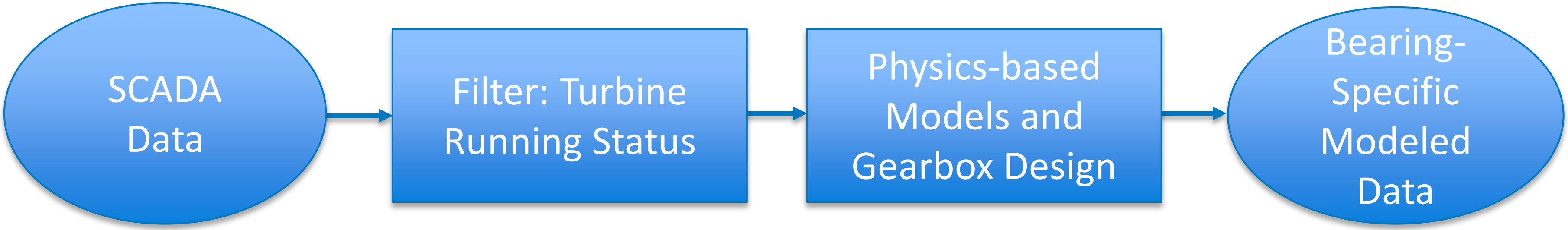
Case Study 1: SCADA and Modeled Data

SCADA Data

- Turbine SCADA data collected from December 2008 to October 2018
- Total: Thirteen 1.5-megawatt wind turbines
  – Axial cracking in either rotor-side or generator-side bearing of intermediate- or high-speed-stage shaft
- SCADA data : 10-minute averaged measurements of various sensors
  – Power, wind speed, bearing temperature, oil temperature, ambient temperature, nacelle temperature, status code, and so on
- A total of 144 (6 per hour × 24 hours) rows of data were recorded per day by a single turbine in the SCADA system.

Modeled Data

- To represent a bearing's health, additional data are calculated using various physics-domain models based on the gearbox configuration information
  - Bearing load, roller load, roller deflection, frictional energy, slide-to-roll ratio, and so on.

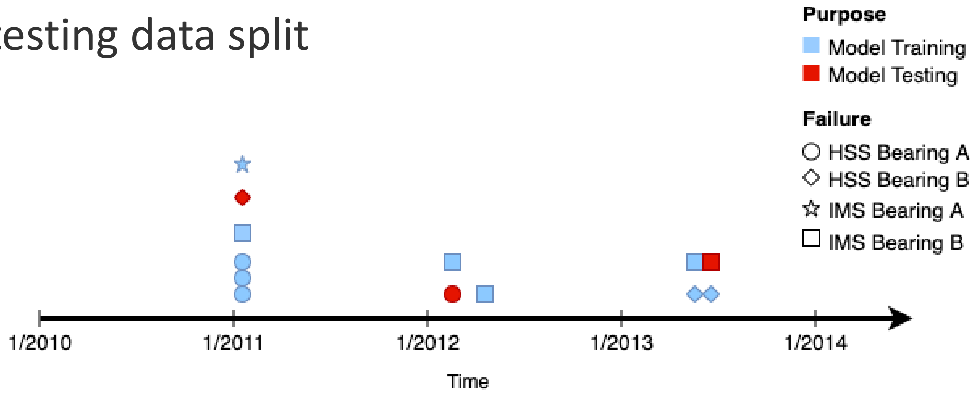


Data Preprocessing

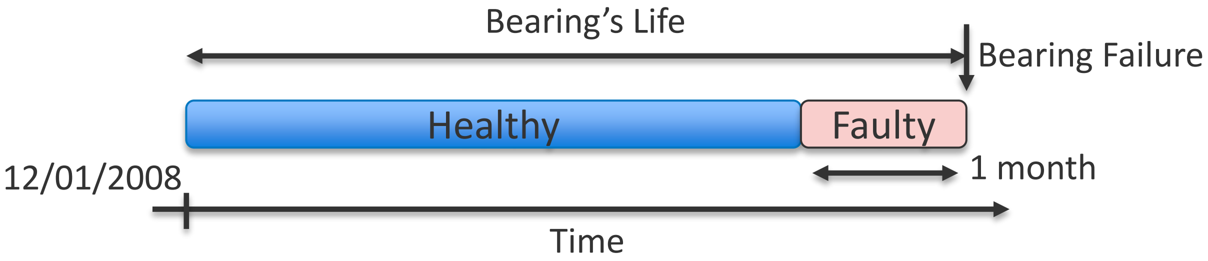
- Consider the data only when a turbine is in running condition and power is produced
- Aggregate data and determine daily summary statistics:
  - minimum, maximum, length of data
  - mean, standard deviation, root mean square
  - skewness and kurtosis.

Data Split and Labeling

- Training and testing data split



- Healthy and faulty data labeling
  - Last month leading to failure labeled as "Faulty"; older as "Healthy."



Class Imbalance Between Faulty and Healthy

- Only last month of data before failure are considered faulty, whereas the majority of the data are considered healthy
  - Problem: algorithms are biased toward majority class
- Two techniques to address class imbalance:
  - SMOTE (synthetic minority oversampling technique)[5]: an oversampling technique generating new synthetic samples in the minority class
  - Cost-sensitive learning[6]: higher penalty for minority class misidentification.

Modeling

- A total of four algorithms are selected for bearing failure prediction:
  - Logistic regression (LR)[7]
  - Random forest (RF) [8]
  - XGBoost (extreme gradient boosting)[9]
  - LSTM (long short-term memory networks)[10]
- Recent historical data are very helpful for prognostics, as axial cracking in bearings does not happen instantaneously
  - Add 1 to 30 days of lagging variables of SCADA channels while training LR, RF, XGBoost
  - LSTM inherently handles a sequence of past observations as input.

Modeling (continued)

- All four algorithms are trained and tested on two sets of data:
  - SCADA data
  - SCADA data and modeled data
- Best or recommended practices adopted in model development
  - Logistics regression, random forest: scikit-learn library[11]
  - XGBoost: XGBoost library[12]
  - LSTM: Keras library.[13]

Performance Evaluation

- Model performance evaluated using Precision, Recall, F1 score, and AUC (Area under ROC [Receiver Operating Characteristic] curve)
- ROC[14] is a probabilistic curve that shows the model performance at various classification thresholds.

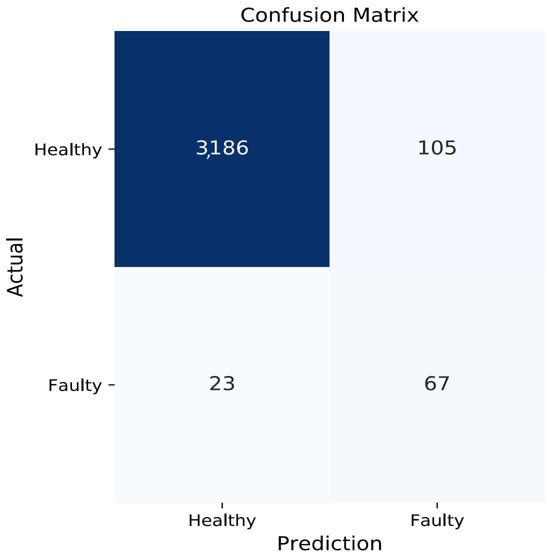
		Prediction	
		Healthy	**Faulty**
Actual	**Healthy**	True Negatives (TN)	False Positives (FP)
	Faulty	False Negatives (FN)	True Positives (TP)

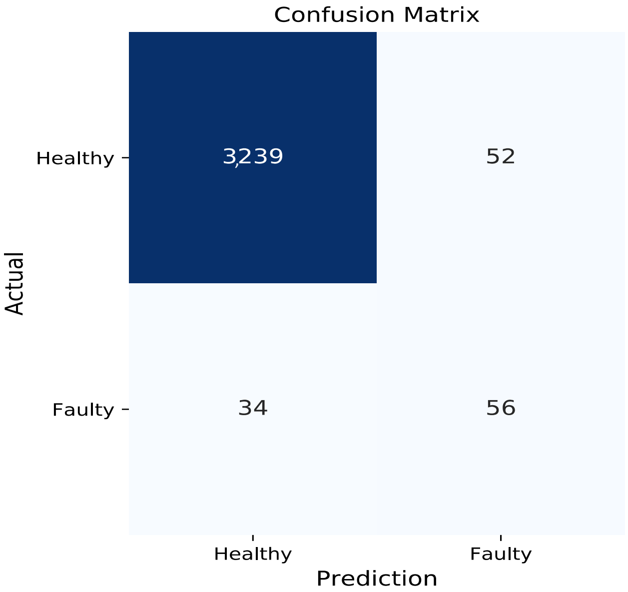
Metric	Formula
Precision	TP/(TP + FP)
Recall	TP/(TP + FN)
F1 Score	2 × Precision × Recall/(Precision + Recall)

Results

- Cost-sensitive learning performs slightly better than SMOTE for logistic regression, random forest, and XGBoost models
- When modeled data are added into SCADA data, model performance increases in terms of F1 score and AUC
  - Precision increases (mainly attributed to false alarms decreases), which improves overall F1 score because it is a harmonic mean of precision and recall
  - LSTM models perform the best (which is not surprising, as they are state-of-the-art algorithms for time-series classification).

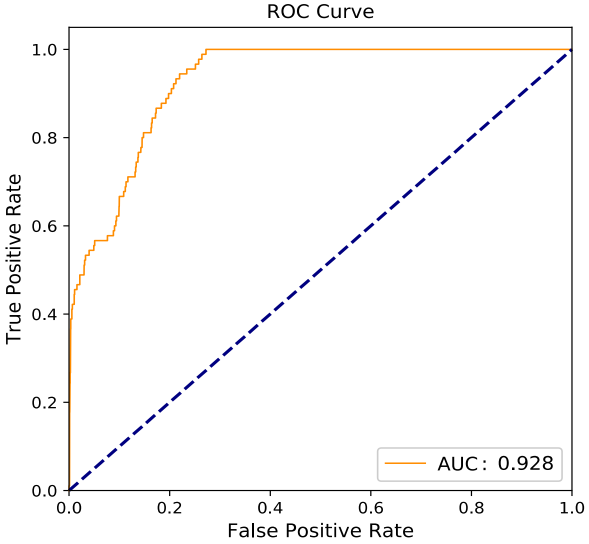
Results (continued)

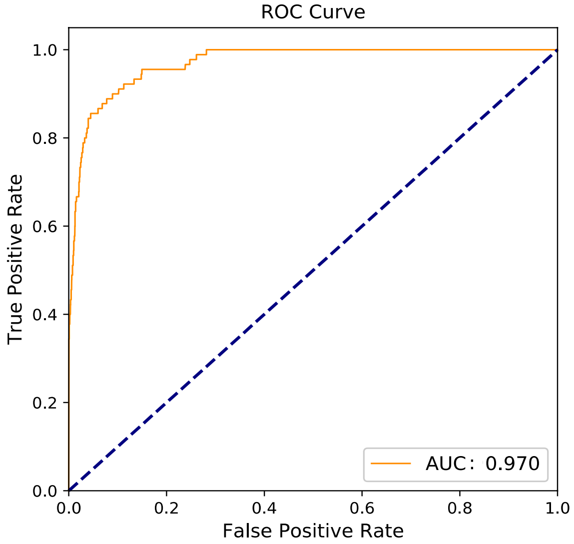




Performance of LSTM model: SCADA data (left), SCADA and modeled data (right)

Results (continued)





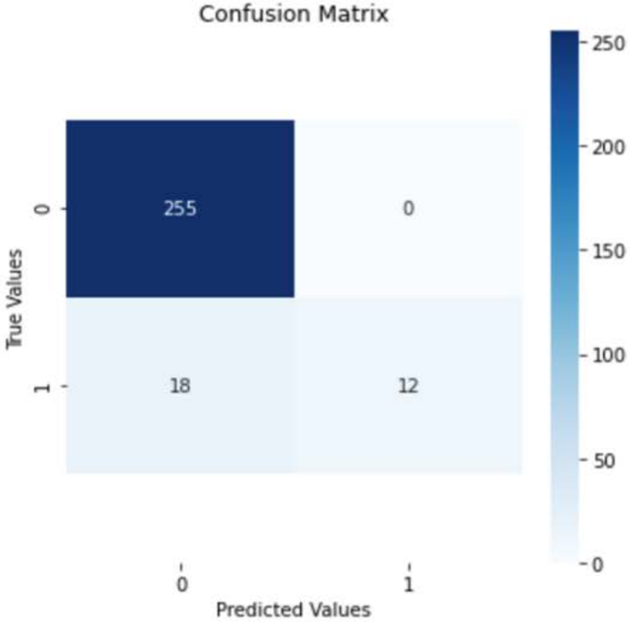
Performance of LSTM model: SCADA data (left), SCADA and modeled data (right)

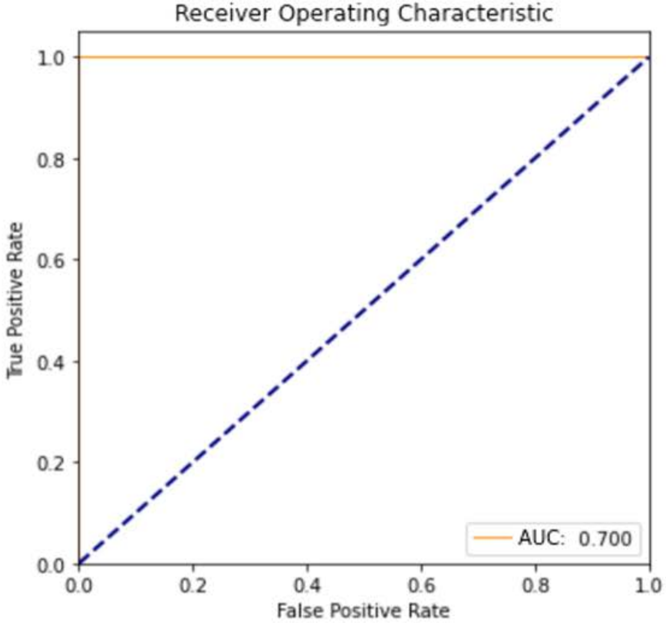
Case Study 2: SCADA, Modeled Data, and Condition-Monitoring Data

Data and Modeling

- A testing site in North America comprising five 2-megawatt wind turbines with identical gearbox configuration, one planetary stage, and one parallel stage
  - Three gearboxes failed by axial cracking on rotor-side bearings in parallel stage
- Available data:
  - Gearbox design, bearing catalog information, and failure records
  - SCADA data, including both time-series and status codes, collected from January 1, 2012, to October 31, 2020
  - Modeled data generated during the same time frame as SCADA
  - Condition-monitoring data
    - Limited vibration data are not sufficient to meet the fault-prediction research needs
    - Oil debris data—specifically, particle generation rates—used as additional input to the machine-learning models as representative of condition-monitoring data stream
  - All streams of data from the three gearboxes with bearing axial cracking are considered up to the appearance of first failure and synchronized to the oil debris data, which has the shortest length
- Data from two gearboxes used for training, and data from third gearbox used for testing
  - Last 30 days before failure labeled as “faulty” and older labeled as “healthy”
- Adopted LSTM, as it was the best performer from pervious research
  - Cost-sensitive learning to handle imbalance between healthy and faulty.

Preliminary Results





- All Data Streams Combined Not Necessarily the Best in Terms of All Metrics
  - Best performance in terms of true negatives, false negatives
  - Not the best in terms of other metrics, such as true positives, false positives, accuracy.

Summary

Observations and Future Work

- Case Study 1 shows the potential of bearing-specific modeled data; the data should be explored further
  - Overall ML model-based gearbox bearing fault prediction can be improved if models for individual bearings are developed
- Case Study 2 demonstrates more data streams may not lead to the best performance
  - Appropriate feature engineering may help
- Optimum classification threshold can be optimized using ROC curve, considering false alarms and cost of missed alarms
- Current research assumes cracks are unknown until they can be visually detected
  - Further work would find optimal time frame (assumed to be the last 30 days in this study) to predict bearing axial cracking.

References

1. Dao, C., B. Kazemtabrizi, and C. Crabtree. 2019. "Wind Turbine Reliability Data Review and Impacts on Levelised Cost of Energy." *Wind Energy* 22(12): 1848–1871. doi: 10.1002/we.2404.
2. Sheng, S. and Y. Guo. 2019. "A Prognostics and Health Management Framework for Wind." *Proceedings of ASME Turbo Expo 2019*. June 17–21, 2019. Phoenix, Arizona: Paper no. GT2019-91533. https://doi.org/10.1115/GT2019-91533.
3. Desai, A., Y. Guo, S. Sheng, C. Phillips, and L. Williams. 2020. "Prognosis of Wind Turbine Gearbox Bearing Failures Using SCADA and Modeled Data." *Proceedings of the Annual Conference of the PHM Society 2020* 12(1): 10. doi: 10.36001/phmconf.2020.v12i1.1292.
4. Guo, Y. S. Sheng, C. Phillips, J. Keller, P. Veers, and L. Williams. 2020. "A Methodology for Reliability Assessment and Prognosis of Bearing Axial Cracking in Wind Turbine Gearboxes." *Renewable and Sustainable Energy Reviews* 127: 109888. doi: 10.1016/j.rser.2020.109888.
5. Chawla, N.V., K.W. Bowyer, L.O. Hall, and W.P. Kegelmeyer. 2002. "SMOTE: Synthetic Minority Over-Sampling Technique." *Journal of Artificial Intelligence Research* 16: 321–357. doi: 10.1613/jair.953.
6. Zadrozny, B., J. Langford, and N. Abe. 2003. "Cost-Sensitive Learning by Cost-Proportionate Example Weighting." Presented at: Third IEEE International Conference on Data Mining, Nov. 22, 2003, Melbourne, FL. doi: 10.1109/ICDM.2003.1250950.
7. McCullagh, P. and J. Nelder. 1989. *Generalized Linear Models, 2nd ed.* London: Chapman & Hall.
8. Liaw, A. and M. Wiener. 2002. "Classification and Regression by Random Forest." *R News* 2/3: pp. 18–22. Available at: https://cogns.northwestern.edu/cbmg/LiawAndWiener2002.pdf.
9. Chen, T. and G. Carlos. 2016. "XGBoost: A Scalable Tree Boosting System." *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 785–794. doi: 10.1145/2939672.2939785.
10. Hochreiter, S. and J. Schmidhuber. 1997. "Long Short-Term Memory." *Neural Computation* 9(8): 1735–1780. doi: 10.1162/neco.1997.9.8.1735.
11. Scikit-learn. (n.d.). "Scikit-learn: Machine Learning in Python." http://scikit- learn.org/stable.

References (continued)

12. XGBoost. (n.d.). “Scalable and Flexible Gradient Boosting.” https://xgboost.ai.
13. Keras. (n.d.) “Keras.” https://keras.io.
14. Fawcett, T. 2004. “ROC Graphs: Notes and Practical Considerations for Researchers.” *Pattern Recognition Letters* 31(8): 1–38. Available at: https://www.researchgate.net/publication/284043217_ROC_Graphs_Notes_and_Practical_Considerations_for_Researchers.

Thank you!

shawn.sheng@nrel.gov

303-384-7106

www.nrel.gov

NREL/PR-5000-81428

Data sharing by wind plant owner operator partners is acknowledged.

This work was supported in part by the U.S. Department of Energy (DOE) Office of Science and Office of Workforce Development for Teachers and Scientists under the Science Undergraduate Laboratory Internships Program.

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Wind Energy Technologies Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government.

NREL Transforming ENERGY