Data Driven Predictive Maintenance & Diagnostics for WindPower O&M Optimization

Bedrock Global A.I. estimates a cumulative sales revenue by licensee of $138M and cumulative licensing revenue of $13.86M during the first 10 years of commercialization.

1 Market Opportunity

To enable wind power as an attractive competitor to traditional energy production, the Department of Energy(DoE) industry research from the National Renewable Energy Laboratory sets a target for levelized cost of energy(LCoE) below $100MW/h. Currently the LCoE of offshore production is about $130-$450MW/h[1]. Of this, operations and maintenance(O&M) costs can be greater than 20% of LCoE and as much as 23%[1,2]. A variety of factors such as- travel time to remote locations, variable harsh weather conditions, availability of spare parts on short notice, availability of skilled personnel, specialized equipment, and onsite time of 50-80 hours as compared to 20-40 hours for onshore contribute to offshore O&M costs[2]. In particular, unplanned maintenance contracts are nearly 25% more than planned events[2].

Our predictive tools will reduce repair costs, decrease downtime, and increase system reliability[1]. This is achieved using data-driven machine learning(ML) modeling of wind turbine operation to identify issues in advance of failure. Consider an unplanned gearbox failure in an offshore asset. A specialized heavy lift vessel and a fair weather window must arrive for the repairs to commence in the open ocean- costs are asset downtime, weather wait time, and cost of ship, equipment, and personnel[1]. Research shows that ML tools as we propose can predict fault/failure of components from two to nine months in advance[Bach]. Such lead time allows for scheduling and conversion of unplanned to planned O&M events, or, reduced production for the “unhealthy” asset slowing the time-to-fault until a better repair time[1]. Lastly, predictions let us recommend efficient asset maintenance routes according to need. This streamlines fuel and labor costs.

Our approach calls for engineering of a suitable streaming dataset whose effort constitutes a barrier to market entry. While SCADA data is generally proprietary, we have a letter of support from DoE funded LEEDco Icebreaker Inc. allowing use of data from their Lake Erie offshore project and we will extend/enrich this data as needed. This support allows us to model operational characteristics of each component from the start of operation, “healthiest”, state onward. Furthermore, we recently gained a research partner whose work extends our capabilities. Dr. David Zeng of Case Western Reserve University is working on engineering approaches to predict structural faults in offshore wind turbines. Also the university operates 3 wind turbines: a Northwind 100, Vestas V-27, and Nordex N-54. On agreement, we will have access to historical turbine data. With this data testbed, we will fuse our partners work with our models to create predictive tools for the wind turbine structure itself. This can require adding custom sensors in precise locations beyond what is available. From this we will pinpoint various component signals and characterize them as a component signature. As the dataset grows, our models will only improve their predictions. In parallel, our network of partners should expand so that we will cover many models of wind turbine.

[1] National Offshore Wind Strategy, NREL, Sept. 2016

[2] Global Wind Power Services Market, Forecast to 2025, Frost Sullivan, MCC6-14, Nov. 2017, subscription database

Our market opportunity is driven by two major factors: (1) In 2016 the global market estimates had a compound annual growth rate(CAGR) of 6.7% for approximately \$46.4B[3] in value, extending this 14 years we expect a market value of approximately \$115B. North America represented 17.1% (actual expected 2017-2022 CAGR of 7.8%[3]) to be conservative assume this remains constant, we see a value of approximately \$19.7B. (2) The average age of the North American wind fleet will be 7 years in 2020 and 14 years by 2030[4]. Newly installed turbines have estimated O&M costs of 15%-20%. This increases to 25%-35% near the end of life cycle[2]. As assets age out of warranty, O&M costs are pushing operators towards performance optimization and cost management[4]. The expected strong growth for the industry and servicing of aging assets can create component supply shortages as the lucrative North American market expands[3]. This may also be relieved to some extent by the lead time afforded by our tools. These macro forces are driving long term market opportunity.

Domestically, we would approach wind farm operators such as Xcel Energy or Avangrid for their expanding projects, beyond this there is a strong backlog of capacity adding projects in planning. It is also true that while our products are meant for the demanding offshore market, they can perform in the less demanding onshore environment domestically and internationally as well.

2 Intellectual Property

Having conducted patent research through both the websearch and the USPTO we examined all patents under CPC subjects: Y02E 10/70, Y02E 10/766, G05B 13/00, G05B 19/4063, G05B 2219/2619, G06N, G06N 5/046, G06N 3/02, and F03D 7/0546. Many patents were a combination of hardware sensors combined with a software solution aimed at diagnostics, not predictive diagnostics in particular. It is conceivable that these companies are prospective customers. Predictive diagnostics are enhancements to their products. For example, G.E.'s Predix system facilitates the "plugging in" of machine learning tools. In other enhancement cases, standard diagnostic snapshots from a suite of sensors are offered without predictive diagnostics/fault detection.

Furthermore, we find that our reservoir computing approach is by no means a commonplace solution. The autoencoder approach for extrapolating possible fault types from a small dataset known faults to infer the space of faults is completely novel. As per the recent development of our partnership discussion with a research institution, any structural fault prediction tool we might develop is an innovation. Any of these are potential candidates for future patenting.

3 Company/Team

Bedrock Global LLC was formed in January of 2019 by a group of three M.S. and one PhD graduate students from the University of California at San Diego(UCSD). All shared a combined fascination for the power of mathematics and the machine learning(ML)/artificial intelligence(AI) algorithms it admits. Our goal is to translate this fascination to meaningful real world AI solutions in novel ways. To illustrate, anyone can use a paint brush but not many can paint as a master. We see a huge opportunity to grow the wind energy market with creative use of the best AI tools available for the job.

Brian Whiteaker is the CEO for Bedrock Global LLC. Prior to returning to school at 38 years old for his B.S. in applied maths and his M.S. in computational science he gained skills in management. As Head Bartender for Claim Jumper Restaurants he was involved in the hiring and training of new

[3] Wind Turbines: Technologies, Applications and Global Markets, BCC Research, GY058C, January 2018, subscription database

[4] US wind OM costs estimated at \$48,000/MW; Falling costs create new industrial uses: IEA", NewEnergyUpdate, 11/22/2017, http://newenergyupdate.com/wind-energy-update/us-wind-om-costs-estimated-48000mw-falling-costs-create-new-industrial-uses-iea

bar staff for multiple stores grossing \$15M/yr-\$20M/yr or more, overseeing that all construction was completed, and liquor laws are observed by trainees. Later as owner of Blackline Training Center in Carlsbad California he was responsible for the day-to-day management of a busy fight gym. Simultaneously, he was head coach and involved in the management of professional mixed martial arts fighters from amateur to the top professional levels of the sport. This experience will be used for business administrative duties at the company. He will also be responsible for strategy and building industry networks. As a data scientist Mr. Whiteaker benefited from his LAUNCH scholarship in data mining which combined with a course load in applied math and ML. Courses in ML required work with deep learning for pattern recognition, computer vision, reinforcement learning, text mining, high performance computing, and other supervised/unsupervised learners. He had internships successfully implementing AI in real world situations. He built a Multi-Armed Bandit for 85SIXTY Marketing firm, then at NASA Glenn Research Center in Cleveland working on the Biomimicry project he built a textmining pipeline with a web-scraper, Latent Dirichlet Allocation and Non-negative Matrix Factorization topic modeling tool, and an ensemble classification tool for use in the project. He was asked back to work on deep neural networks to predict jet engine acoustic vibration noise on aircraft surfaces. For this his graduation is postponed until June 2019.

Jeremy Schmitt is a PhD in mathematics, M.A. in applied Mathematics. With a heavy background in numerical analysis- his thesis and research papers are in the area of Hamiltonian variational integrators. He received an award for excellence in teaching, and prior to his defense was refereeing the Elsevier Journal of Applied Mathematical Modeling. At FICO he is a member of their research staff and specializes in fraud detection. He received the FICO Spot award for his machine learning work on their Falcon Fraud model. Jeremy provides great depth to the development process at the Bedrock.

William Mendoza-Gopar is an M.S. in computation from the Computational Science Mathematics and Engineering program(CSME). He is a LAUNCH scholarship recipient and also has a deep applied maths background. Mr. Mendoza-Gopar has worked as an intern at Autodesk where he constructed a pipeline for their streaming server data. He was tasked with constructing an automated system to clean and analyze Terabyte sized datasets in real-time.

Premanand Kumar was a working mechanical engineer before returning to school for his M.S. in computational science from the CSME. He successfully developed a rule based system for his company's manufacturing processes. Since receiving his M.S. he worked for Locbit and developed an ML prediction algorithm for the energy consumption of buildings. His work is going into production for Locbit.

As of December 2018, Bedrock Global LLC began work for its first client Field Intelligence Incorporated. Field Intelligence supplies diagnostics on globally streaming data for fleets of diesel powered centrifugal pumps used in mining and agriculture. Bedrock is developing ML fault prediction tools for the pumps to recognize such faults as cavitation, rotating cavitation, bearing failure, axle mis-alignments and others. This effort dovetails well with our work on topic 15b. We anticipate that our revenue from the concurrent Field Intelligence work and our SBIR Grant is sufficient to sustain our Phase 1 development effort.

4 Revenue

For our business there are two macro forces leading to revenue: (1)Expected growth in capacity (2) Expected growth in operations and maintenance(O&M) needs for aging assets. Extending our reach further is the adaptability of our tools for onshore wind and international markets. It is noteworthy that there are many projects planned for offshore production in the U.S. market, the outlook is for compound annual growth rate(CAGR) of 39.4% from 2016 to 2022 in this space[3]. We will calculate projections somewhat conservatively though.

The global market in 2016 was \$46.4B and is projected to be \$71.2B in 2022[5]. BCC Research reported the North American wind turbine market was valued at $9.3B in 2016[5]. The expected CAGR from 2016 to 2022 is 7.8%. If we extrapolate these values through for North America to 2030 we find an expected market of $26.6B. This is a $17.3B increase over the time span, which at constant rate is $1.24B/year. Using this number for ease we see North American share being $12.4B over the next 10 years.

As a possible revenue example for Bedrock Global, suppose we add a competitive advantage to Suzlon's wind turbines. As of 2017 they accounted for 2.9% of world market share. Using this value for North American share and assuming this remained the same through 2030, 2.9% of $12.4B, then Suzlon revenues would be $359.6M in North America. Supposing we license for 10% and we are responsible for 30% of sales. This is $107.9M which we charge 10% in licensing on. then we have 10 year revenues of about $10.8M when supporting OEM growth in North America.

As wind assets age out of warranty more operators wish to house their own servicing. To fill this gap, Original Equipment Manufacturer's (OEM) and Independent Service Provider's (ISP) are moving to provided cross-manufacturer servicing capabilities. North American services market revenue grew by 10.5% in 2016 to reach $1.8B and forecast a CAGR of 9.7%[2]. As above we extrapolate the market to 2030 for a market revenue of about $6.58B. As of now there is a race to market in this field, who becomes dominant is unclear, however many ISP's can gain a competitive advantage by licensing our service. Assume that the increase from $1.8B to $6.58B of $4.78B was spread out equally, this is $341.4M/year. Over 10 of 14 years this market value is $3.4B. If a company maintained constant 3% market share this is $102M. Assuming we account for 30% of revenue and licensing our tools at 10%. We generate $30.6M in revenue for the licensee and charge for $3.06M license fees.

So our combined 10 year revenue generated by the licensee's is about$138M and our revenue from licensing fees is $13.86M.

[5] "Wind Turbines: Technologies, Applications and Global Markets", BCC Research, GY058C, January 2018, subscription database

Brian Whiteaker, Bedrock Global A.I.
15B: Data Driven Predictive Maintenance & Diagnostics for Wind Power O&M Optimization

Project Summary/Abstract

The Department of Energy National Offshore Wind Strategy and Wind Vision have laid out actionable issues to position off-shore wind power as a more competitive clean source of energy. Reduction of the levelized cost of energy from \$130 - \$450 MW/h to under \$100 MW/h is desired through overall reductions in operation and maintenance costs which accounts for 24% of the total levelized cost of energy.

Bedrock Global A.I. is addressing this problem by reducing operation and maintenance costs through data-driven, predictive maintenance technology. We will use state-of-the-art machine learning methods to alert in advance of critical failures by making use of data emitted through sensors directly placed on the wind turbine. Significant cost reductions result from successful predictive maintenance strategies by reducing unplanned maintenance and downtime. This reduction in operation and maintenance costs will help reach DOE target levelized cost of energy for off-shore wind turbines to be under \$100 MW/h.

During the first phase we be working closely with our wind turbine partners to receive sensor-data necessary to begin the development of our predictive maintenance solution. Our team of data scientists will be constructing machine learning algorithms for early fault detection in a wind turbine. Our systematic approach will be constructed by several modules which include: data preprocessing, data augmentation, a core machine learning engine, and a visualization engine. Our solution will be tested in a real wind farm setting by the end of phase 1.

Aside from making wind-farms far more efficient in energy production by increasing system reliability, the development of this project will encourage the use of domestic sources of renewable energies that prevent emissions of environmental pollutants and greenhouse gasses. In addition, there is an expected market of \$70B/year indicating massive scale wind-farm operations that will potentially benefit from our project, and also create a new job sector in wind-turbine maintenance and increasing the number of on-site technicians.

Key Words: Machine Learning, Artificial Intelligence, Predictive Maintenance, Fault Detection, Offshore Wind Energy, Levelized Cost of Energy.

Summary for Members of Congress: Reductions in Operation and Maintenance (O&M) costs for off-shore wind power is needed to position it as a more competitive and desired source of energy. Predictive maintenance of turbine components can be realized using state-of-the-art machine learning methods using sensors mounted to the turbine to prevent unexpected down-time increasing reliability while driving down O&M costs.

Project Narrative

Data Driven Predictive Maintenance & Diagnostics for Wind Power O&M Optimization
15B

Bedrock Global, LLC
PI: Brian Whiteaker
1 Miramar St. 929478
La Jolla, CA 92092-0003

1 Proprietary Data Legend

N/A

2 Identification and Significance of the Problem and Technical Approach

The Department of Energy (DOE) seeks to increase the viability of offshore wind power as a cost effective alternative production source. With the global wind power market expanding at a cumulative average growth rate(CAGR) of 6.7% to an estimated $71.2B by 2022 reports such as the National Offshore Wind Strategy recommend a levelized cost of energy (LCOE) below $100MW/h from the current range of $130-$450MW/h[1], is optimal and achievable. We focus reducing the Operations and Maintenance (O&M) cost which is approximately 23% of the total cost[1]. The environment for offshore wind is harsh and dynamic and presents logistical challenges. Any unplanned O&M event requires specialized ships with heavy equipment operating in a window of fair weather conditions, onsite for an extended time. Hence, the estimated costs of unplanned events can be as high as 25% more than the planned cost[2]. Hence, we are building a solution to reduce unplanned events by machine learning/A.I. techniques. The state-of-the-art machine learning tools we propose to build, train, and continually validate during wind farm operations are driven by real-time streaming sensor data. They will provide predictive diagnostics and fault prediction to the owner. This lead time allows for scheduling of maintenance assets resulting in cost control/reductions, reduced downtime, reduced hazard to onsite personnel, and more efficient allocation of O&M assets.

2.1 Overall Technical Approach

The technical problem to be solved is to accurately detect and identify an evolving fault state in an offshore wind turbine (OWT). We use the term *fault* to mean the abnormal functioning or degradation of a wind turbine component. *Failure* is the cessation of function of a wind turbine component or for the wind turbine as a whole. A prediction of the wind turbine future state, for use in our predictive maintenance routine, involves the following: dataset engineering, robust model construction, and model enrichment through known engineering equations. Dataset engineering and robust model construction aim to generate a model achieving a high true-positive ratio, a low false-positive ratio, and providing significant lead time. An accurate prediction is useless when the lead time is too short, and likewise, a large lead time is useless without a high-level of accuracy in the prediction itself. A prevalence of false positives diminishes any efficiency gains coming from predictive maintenance. It is significant to modeling that failure states (abnormal operation) are much less common than non-failure states (normal operation). Thus the problem of failure prediction is similar to the field of anomaly detection, and the technical problem can be re framed as one of density estimation. While sensor data from the wind turbine supervisory control and data acquisition system (SCADA) data provides a homogeneous data format, the statistical nature of wind turbine datasets is often heterogeneous. This may be attributed to the geographic location of the turbine, seasonal variances of said location, maintenance practices, and differences of turbine design. Heterogeneity can impede standard statistical and machine learning(ML) methods. For effective model generalization special methods must be used to counter heterogeneity. In engineering our dataset we will identify informative features and design more powerful features by capitalizing on the time-series nature of SCADA data. Methods such as vibration analysis will uncover dependencies and important signals in the data, which we will separate out for better modeling. To enrich the dataset and, in turn our

[1] National Offshore Wind Strategy, NREL, Sept. 2016

[2] Global Wind Power Services Market, Forecast to 2025, Frost Sullivan, MCC6-14, Nov. 2017, subscription database

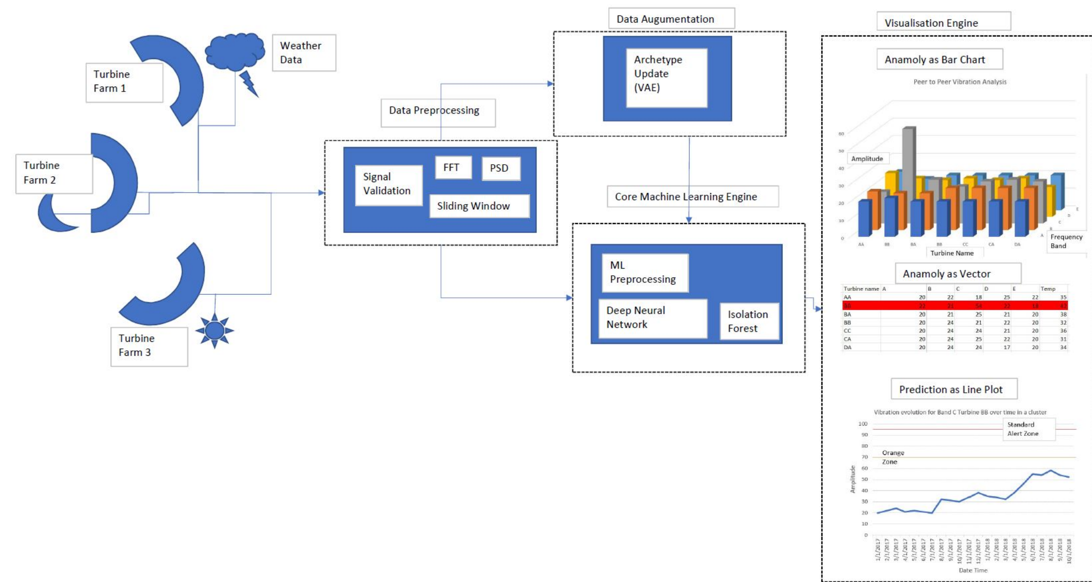

Figure 1: Predictive Maintenance Methodology

models, we may rely on the extensive engineering that has come before. The design and engineering of wind turbines relies upon equations describing the mechanics of the dynamic physical system, it is wasteful to ignore their value when implementing a predictive maintenance routine. We believe classical approaches can pair with more recent ML approaches in a complementary fashion. A simple method we will utilize would integrates the known equations describing a wind turbine component as an input to the dataset. We will also use them as a constraint to the validity of the models. Work in this field is still in its infancy and so we will be expanding on recent advances as a matter of course.

2.2 Outline of the proposed solution

Given the inherently complex dynamics of the offshore wind turbine system, machine learning solutions offer many different approaches that can be combined in solving the underlying technical problem. No one method can lay claim to solving the problem of predicting complex/chaotic dynamical systems, and correspondingly, we will utilize a multi-pronged approach using both unsupervised and supervised techniques as well as some standard tools from time series analysis and signal processing.

The key components of our predictive maintenance solution are the following: (1) sensor data from the wind turbine supervisory control and data acquisition system (SCADA) and weather data (2) signal validation and conditioning (3) data transformation/pre-processing and feature engineering (4) Variational Autoencoder (5) Online Archetype Soft Clustering Unit (6) Deep Neural Network (7) Iso Forest (8) Visualization Engine. Figure 1 displays how the different parts of our predictive maintenance methodology will integrate into a real-time fault detection system. SCADA and weather data transformation/pre-processing will involve sliding windows and time-decayed values to capture key dynamic aspects of the time-series data. A variational autoencoder (VAE) trained on normal/fault-free data will create archetypes of normal OWT function/behavior. Via the VAE, each OWT will have a distribution over these archetypes that will be updated dynamically in real-time with a novel low-memory high-throughput algorithm, thereby allowing online detection of behavior deviation based upon the current state of the OWT and the states of its peers (as determined by the archetype distribution). The recent states of the OWT will then determine the feed-forward neural network that will take the transformed data and archetype information as input and produce the probability of upcoming fault/failure.

2.3 Research portion of our solution

Since, the data streaming and visualization techniques are more of an engineering challenge, the research during phase 1 will be to provide machine learning models to the core machine learning engine and data augmentation processes that will enable accurate prediction of failure. Enabling a connection from the core machine learning engine to a highly interactive dashboard can then be created for real-time monitoring and inspection through easy to use user-interfaces.

2.4 Contemporary Technical Solution

GE , Mitsubishi Vestas Offshore (MVOW), Wind and Siemens are major original equipment manufacturers (OEM) with predictive analytics tools offering practical solutions and robust use-cases. Our technical solution will consider heterogeneity, normalizing the operation and weather conditions, to get meaningful comparison with a baseline value is the key challenge to any diagnosis. MVOW combats this in one of its patents for vibration analysis by creating a ‘frequency index’ which is normalized by the rotating speed of the shaft thereby filtering the operating load. And instead of filters on the weather condition, they choose to compare operation within the farm establishing that the weather condition is the same for all the turbines located in proximity[3]. Thus paving the way for a peer to peer real time comparison within the wind farm. Siemens employs Neural Networks with weather features such as wind and temperature as inputs and also includes historic behaviour of the turbine to calculate the optimal run time settings (e.g rotor blade angle)[4]. In addition, one patents discusses the use of Recurrent Neural Network to determine the future states[5]. It then determines the specific action to be set for optimal operation as the system learns from historic data.

3 Anticipated Public Benefits

3.1 Technical Benefit

Machine learning technology is a fast growing component of the Industrial Internet of Things(IoT). Joshua Bloom of GE, the global leader in this space, acknowledges that domain expertise is required for successful ML integration to industrial world[6]. This poses a different challenge when compared to traditional ML applications in the consumer world. Thus, our proposed project will be a technological feat in adopting known ML techniques to a new frontier in wind power generation and hence in the industrial world in general. Furthermore, successful delivery of data driven predictive maintenance technology for reductions in O&M costs will allow domestic offshore wind power to be a more desired and reliable source of renewable energy while also pushing boundaries in the IoT sector.

[3] Lim; Knoon Peng, Zhou; Yu, Chen; Wanying. Method for performing condition monitoring in a wind farm, US 20130261988 A1, USPTO, July 24, 2018

[4] Ulrich Kreutzer, Machine Learning: Optimizing How Wind Turbines Work (article)

[5] Schafer; Anton Maximilian, Udluft; Steffen, Zimmermann; Hans-Georg. Method for the computer-assisted control and/or regulation of a technical system where the dynamic behavior of the technical system is modeled using a recurrent neural network, US 20100094788 A1, USPTO, October 8, 2013

[6] Joshua Bloom, Strata Data Conference in Singapore, 2017

3.2 Economic Benefit

The increases of wind farm operations nationwide provide local economic development opportunities for rural project host communities as well as equipment and hardware manufacturers around the country[7]. Up through 2014, the wind supply chain supported well-paying manufacturing jobs at over 500 domestic manufacturing facilities in 43 states, totalling an average of 73,000 total jobs in the years of 2010-2015[8]. It has also been shown that wind power projects installed between 2000 and 2008 found an aggregate increase in county-level personal income and employment of approximately $11,000/MW and 0.5 jobs per megawatt of wind power capacity installed[9]. Wind power increases domestic fuel diversity, making the overall electric sector 20% less sensitive to changes in fossil fuel costs. Long-term costs of wind power create downward price pressure on fossil fuels that can cumulative save consumers $280 billion from lower natural gas prices outside of the electric sector[10]. Our data driven predictive maintenance solutions built to drive wind power O&M costs down, has strong potential to accelerate and facilitate these public benefits in an offshore setting.

3.3 Social Benefits

Literature has also demonstrated significant societal impacts from wind power. On a life cycle basis, wind power has among the lowest levels of greenhouse gas emissions over different energy technologies, and reported an estimated reduction of 115,000,000 metric tonnes of carbon dioxide and 157,000 metric tonnes of reductions in Sulfur Dioxide in 2013[10]. The technological, economic, and social benefits through successful predictive maintenance technology is substantial.

3.4 Beneficiary groups

Direct customers who benefit from our project are Siemens, MHI Vestas, Suzlon, Aegis Wind and other individual turbine owners with out of warranty turbines. DoE will indirectly benefit from our effort to reduce the LCOE cost.

3.5 Resultant product

The end product would come as a software component continuously monitoring data generated through vibration and temperature sensors that reduce repair costs, downtime through software predictions of catastrophic component failures. The expected strong growth for the industry and servicing of aging assets can create component supply shortages as the lucrative North American market expands. A nationwide need for robust predictive maintenance technology will produce markets for our product.

4 Technical Objectives

1. Interpret the heterogeneity of the data with respect to different operating and weather conditions. Create homogeneous archetypes/clusters to enable peer-peer comparison across wind turbines in the network.

2. Reconcile availability limited of fault/anomalous data. Use homogeneous clusters and synthetic data to combat the relative lack of fault/anomalous data.

[7] Jose, Zayas, et al. "Enabling Wind Power Nationwide." 2015

[8] American wind power rebounded in 2014, adding over four times as much as year before (press release). Washington, D.C.: American Wind Energy Association, January 28, 2015.

[9] Brown, Jason P.; Pender, John; Wiser, Ryan; Lantz, Eric; and Hoen, Ben, "Ex post analysis of economic impacts from wind power development in U.S countries" (2012). Publications from USDA-ARS / UNL faculty. 1144. http://digitalcommons.unl.edu/usdaarsfacpub/1144/

[10] DOE 2015. https://www.energy.gov/eere/wind/maps/wind-vision

3. Create ML models to detect faults with considerable lead time, a low false positive ratio and a high true positive ratio. This will be the heart of our core ML engine.

4. Create a complete software module that combines all the above three technical objective and produce a predictive solution as depicted in Figure 1.

5 Work Plan

Phase 1 will be used to produce working prototypes of ML models working with streams of sensor data from wind turbines located in multiple geographical locations. The team has gained support from Case Western Reserve University and LEEDCo to provide the required turbine data.

5.1 Task 1: Prepare Data Pipeline

As depicted in Figure 1, we begin by preparing a data pipeline for our core machine learning engine. To do this we will utilize the Python programming language and the open-source NumPy and Pandas libraries. These libraries are a basis for developing custom data processing and transformation modules assisting wind turbine specific model prototyping, data exploration, and visualization. Regular interaction with domain experts for data validation are a must and our team members with expertise in data processing will be vital. As a company, we have recent experience in this process with Field Intelligence whose global pump installations required substantial amounts of data processing. For data visualization work, our team will use both Plotly and Bokeh visualization libraries.

5.2 Task 2: Create specialized tools for data preprocessing.

We will create custom tools required for data validation and preprocessing such as sliding window, time-decayed averages, Fast Fourier Transform etc. Team members will provide basic plans for the design and capabilities of our specialized library modules. Technical details regarding input data formats needed for specific machine learning models will be identified. Engineered features for non-linear, non-stationary time series will be constructed as needed.

Variable selection and feature engineering are critical for any successful predictive model. Some of the variables to be used include those for bearing temperature and vibration, which have been demonstrably useful in fault prediction of the OWT gearbox and bearings [7], [1], [2]. Of the variety of faults that can occur in a wind turbine system, a survey of wind turbine fault studies noted that electrical and electronic control system failures were the most common, but gearbox and generator bearing faults/failures led to the longest downtime [8]. Features will be engineered with this in mind. We will be implementing feature engineering based on the use of sliding windows, moving averages with exponential smoothing, and auto-regressive formulas that provide the network with powerful information about the past and present state of the time-series data. Our time-decayed averages with exponential smoothing are based on the following standard formula,

$$s_0 = x_0 \tag{1}$$

$$s_t = x_t + (1-\alpha)s_{t-1} \tag{2}$$

where $t > 0$ and $0 < \alpha < 1$. The rate of time decay is determined by choosing τ (i.e. $\tau = 30$minutes) and setting the smoothing parameter $\alpha = 1 - e^{-\Delta T/\tau}$, ΔT is the time between measurements, which averages around 10 minutes for wind turbine SCADA data. This form of a moving average (or alternatively, a low-pass filter) has two benefits, (i) the current value is based on all past values, yet it exponentially decreases the impact of later times giving heavier weight to current values (ii) it only requires that the most recent value be stored, resulting in low memory requirements. Letting

s_t^τ denote the τ- time decayed exponential moving average, then normalized will be created by dividing $s_t^{\tau_1}$ by $s_t^{\tau_2}$, where $\tau_1 < \tau_2$. Additionally, time decayed maximum and minimum values will be created. The previous n states of the OWT will be stored in a compressed form and fed into our model, thereby implying that our model is in the class of nonlinear autoregressive models. A large family of features based on these simple formulas will be used during model development, although not all will be selected for use in our production model. More traditional vibration analysis [7] will also be utilized via data collected from sensors placed throughout the OWT relevant to the bearings and gearbox.

Our approach to vibration analysis is by no means novel, but it will enhance and complement our ML solutions, so we briefly mention it here. Using standard techniques from signal processing involving the Fast Fourier transform, band pass filters, Hilbert transforms, dynamic sampling, and envelope analysis the various frequencies coming from the bearings and gearbox can be isolated and a baseline measurement of the OWT will be established early in its operational life span. The baseline will be specific to each OWT and will depend upon mechanical design elements such as bearing size and gear ratios. Once the baseline or signature of the OWT is established, then online comparison of the current operational vibration signals with the baseline will allow for the identification of faults such as deformations in the bearing runways or broken teeth on the gears. Some of the signal measurements will also be used as features in our neural network model, however, the vibration analysis will also serve as a stand alone system that can be used to isolate areas of fault that are indicated by other parts of our system.

5.3 Task 3: Create Archetypes Model

We will use a combination of scikit-learn, Keras, and PyTorch libraries for prototyping models under this section. The team will be work on creating the soft clustering/archetype models, while also confirming the homogeneity of the clusters and validity of synthetic data. A clear metric on homogeneity will be developed with respect to OWTs to assess our efforts to nullify heterogeneity.

Statistical heterogeneity in a dataset can significantly slow learning and degrade overall performance. Offshore wind turbine (OWT) can have a great deal of heterogeneity due to different environments of operation, time of year/day, weather conditions, age of the OWT, or design of the OWT. One common approach for dealing with statistical heterogeneity is to segment or hard cluster the data via expert rules or by using unsupervised learning algorithms, then each segment/cluster can be trained separately. Alternatively, soft clustering can be used in place of or in unison with hard clustering. The distinction being that hard clustering assigns each data point to exactly one cluster, while soft clustering assigns to each data point (i.e. current state) a distribution over the clusters. We will be using a combination of hard and soft clustering, but it is our approach to utilizing the soft clustering that we believe sets us apart. Using a hierarchical model in our soft clustering approach, each cluster can be viewed as an archetype of OWT behavior, where an archetype is defined as a distribution over states or types of behavior, and each OWT is dynamically given a distribution over archetypes based on past and current behavior. One archetype may represent normal operating states of an OWT during the early morning when the wind is milder or an archetype may represent an older OWT with a particular type of drivetrain, etc. Figure 2 displays this hierarchy.

5.4 Task 4: Create Models for Archetype Update & Data Synthesis

Using a variational autoencoder (VAE) and fault-free data, the creation of fault-free archetypes will be done via variational Bayesian inference during the model development period. Then once in production, using a novel algorithm each OWTs distribution over the archetypes will be updated in real-time and with low-memory requirements, allowing for behavioral changes that may correspond to the time of day, seasonality, component degradation, or the age of the OWT. Let $\theta^{(t)}$ denote the archetype distribution for a given OWT at time t, then $\theta_i^{(t)}$ is the probability of archetype i at time

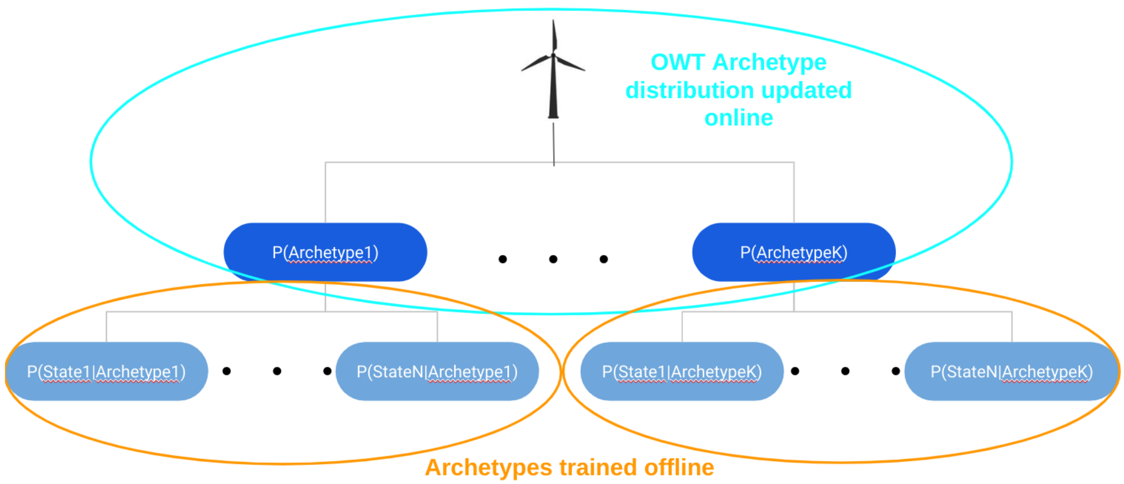

Figure 2: Hierarchy Clustering

t. Given the archetype distribution, the probability of an OWT being in state m at time t is given by the formula

$$P(m|\theta^{(t)}) = \sum_i \theta_i^{(t)} P_{\theta_i}(m) \tag{3}$$

Where $P_{\theta_i}(m)$ is the probability of archetype i being in state m. The upshot is a generative model that provides a measure of "surprise" for the current mode of operation for the OWT. Surprise, defined as $-\log P(m|\theta^{(t)})$, serves as one measure of deviation from normal fault-free behavior. A normalized measure of surprise is obtained by first dividing, $\frac{P(m|\theta^{(t)})}{P(m)}$, where $P(m)$ is the probability of being in state m over all OWTs in the segment. Other measures of deviation will be obtained by storing snapshots of the archetype distribution over time, then comparing the past archetype distribution to the current distribution via Jensen-Shannon divergence and cosine similarity. Any significant deviation from baseline behavior can indicate faults or failures in the near or medium-term. The various measures of deviation provide a measure of confidence and seriousness of the deviation from normal behavior that is obtained not simply by looking at the past history of the specific OWT, but by leveraging the information in the archetypes that is derived from the entire set of OWTs in the segment. Thus surprise carries not only information local to the specific OWT, but also global information from other OWT peers. Updating the current OWT archetype distribution will be done via a form of the following algorithm,

$$\gamma_i^{(t+1)} = P(i|s_{t+1}, \theta^{(t)}) \tag{4}$$

$$\theta^{(t+1)} = \frac{t}{t+1}\theta^{(t)} + \frac{1}{t+1}\gamma_i^{(t+1)} \tag{5}$$

where i represents archetype i and s_{t+1} is the state of the OWT at time $t+1$. Note the distinction between using $P(i|s_{t+1}, \theta^{(t)})$ and $P(i|s_{t+1}, s_t, \ldots, s_0)$ .The later requires the storage of all past states, while the former compresses this information into the archetype distribution vector $\theta^{(t)}$, which allows for a quick and low memory update. It should also be noted that the above update formula will be augmented via exponential time-decay to give more weight to most recent states. Note the measures of surprise and behavior deviation are independent of the faults in the training data, since the archetypes are derived without fault data and trained in an unsupervised manner. This provides additional robustness to unseen faults. The various quantitative measures of surprise and deviation along with the latent archetypes will also be used as features in a supervised feed-forward neural network. The archetypes themselves will also help in determining the origin of the fault, as archetypes are often far easier to interpret than the more opaque output of neural networks. Additionally, the archetypes of an OWT operating state can be used to guide a form of model ensemble. The above approach helps to address not only statistical heterogeneity, but also the broader goal of fault and anomaly detection.

With a similar VAE we can take advantage of the learned latent space in an innovative way. Beyond creating archetypal distributions for fault free operation, we may do the same with any known faults or failures present in the historical data. The benefit is that as we accumulate faults over time the faults will still be anomalous compared to fault-free instances. Can we gain a mapping of the conceivable fault space? The latent space of the VAE provides us with an opportunity to accomplish this. The latent space containing the distributions a VAE has learned from training data may be sampled just as one might sample points from a Gaussian. If we train and generate archetypes on fault/failure data this will begin to form a knowledge of how they combine to form the space of all faults. We may sample from this space to generate "mock" faults and check them against a constraining engineering equation for their viability as a real failure. As the dataset and faults over time naturally increase, we may aggressively model the contributing factors to faults. This will inform routine maintenance of where to pay attention, even when operation is within acceptable bounds, but an aberration is detected.

5.5 Task 5: Create Models for Fault/Anomaly Detection.

We will use sci-kit learn and PyTorch libraries for prototyping ML algorithms and creating custom libraries as required. Our team of data scientists will work on creating the models for anomaly detection. The created models will include VAEs, deep neural networks, and Isolation Forests, to name a few. The task will include identifying possible failure situations and developing metrics to track efficiency degradation over time. While the feed-forward neural network structure is not explicitly aimed at time-series structure in comparison to recurrent neural networks (RNN), our feature engineering and feature selection will contain relevant time-series information of the past and current state. Members of our team have extensive experience in precisely these types of models, and we have them found to be quite successful in other applications of time-series prediction. Other researchers have found similar models useful in the case of OWT prediction [1].

A new type of autoencoder that we will be implementing is known as a meta-learning autoencoder (MLA) [9]. The benefits of this method are three-fold. First, it is geared towards learning varying tasks that have a common underlying mechanism. In this case, we will have a variety of OWTs that all follow a broad mechanical structure (i.e. involving rotational energy and a drive-train). MLA has shown the ability to learn from smaller datasets by learning and benefiting from the underlying mechanism that is common across the data. Secondly, MLA can be used to learn the smaller dataset of true mechanical fault/failure by discovering latent features that can serve as an early warning signs of upcoming fault/failure, and to generate synthetic fault/failure data. Lastly, the likely glut of data on normally operating OWTs may need to be down-sampled in our supervised modeling to deal with class imbalance and to maintain computational efficiency/feasibility. MLA offers a way to determine the information value of each data point that can guide this downsampling process.

Additionally, we will consider an algorithm aimed directly at the problem of anomaly detection that involves a unique twist on decision tree classifiers. This algorithm is known as an Isolation Forest (IF) [4]. IF rests upon two fundamental assumptions, which are satisfied by the very definition of an anomaly (1) anomalies are the minority class, whose number of instances is much smaller relative to the number of normal instances (2) some of the attribute values of an anomaly are very different from a normal instance, and may be considered "incompressible" as opposed to in-liers which may be represented easily by some basis. Based on these assumptions it follows that a binary decision tree would on average "isolate" anomalies much closer (with an appropriate choice of distance or path-length) to the root of the tree, as compared to distance for normal instances. The upshot is an algorithm that has linear complexity, low memory requirements, and a robust training procedure that works well with small or down-sampled data sets.

Very recently, we have found encouraging results from the area of reservoir computing that could be quite applicable to the underlying technical problem of non-linear, non-stationary time series prediction. Due to its cyclic structure, the recurrent neural network (RNN) is naturally suited for time-series data. While the most widely known RNN is the long-short term memory (LSTM) [3], we will explore the potential of a particular type of Echo State Networks (ESN) that belongs to the broader class of Reservoir Computing (RC) [5]. Very recently there have been some startling breakthroughs in predicting chaotic trajectories via the use of RC [6]. In particular, the most successful techniques have partitioned the phase space of the dynamical system, then assigned a RC for each partition. This allows each RC to learn the dynamics particular to a smaller region of the phase space, which is ideal when dealing with the widely-varying dynamical behavior associated with non-linear non-stationary time series. Offshore wind turbines (OWT) have varying levels of operation and environment that correspond to different regions of the phase space. Training separate RCs for each broader mode of operation has the potential to significantly boost the power and performance of our machine learning solution. One concern may be that training multiple RNNs on subsets of the data will be particularly prone to overfitting due to the large number of model parameters involved. However, RC only trains the weights from the hidden to output layer and randomly initializes fixed weights from a sub-Gaussian distribution, where the underlying distribution is chosen based on the properties of the underlying dynamical system. In addition, to solving the problem of vanishing gradients, this method of training significantly reduces the number of parameters to be fit, thus decreasing the likelihood of overfitting.

We will use our collection of deep feed-forward neural networks (DNN), each based on a particular segment or archetype [Figure 3], to form a novel ensemble weighted by the current archetype distribution. Ensembles provide addtional robustness and decrease the chance of overfitting. Features such as raw SCADA input, time-decayed values, sliding windows, vibration frequency analysis, surprise, and archetypes will form the primary inputs into these DNN. The advantages of choosing this form of DNN over a RNN is realized both in terms of more robust model development and better online computational efficiency. Advances in artificial neural networks have shown promising results in the prediction of non-linear, non-stationary time-series. We believe these novel and developing techniques will be a powerful tool in dealing with the non-linear, non-stationary time series associated with OWT data.

5.6 Task 6: Demo the tools on live streaming data

This task is composed of connecting our in-house developed solutions to real live streams of data produced through operational turbines. Our core ML engine will be connected to databases of LEEDco and Case Western Reserve turbines and our diagnosis procedures will follow. This demonstration will also indicate the advantage of having a real-time peer to peer comparisons and also alert on any impeding faults based on our trained core ML engine.

6 Link the Work Plan to the Technical Objectives

Tasks 1 and 2 create the environment for other tasks. They create a platform to feed data into the machine learning engine and data augmentation module. The visualization engine displays the

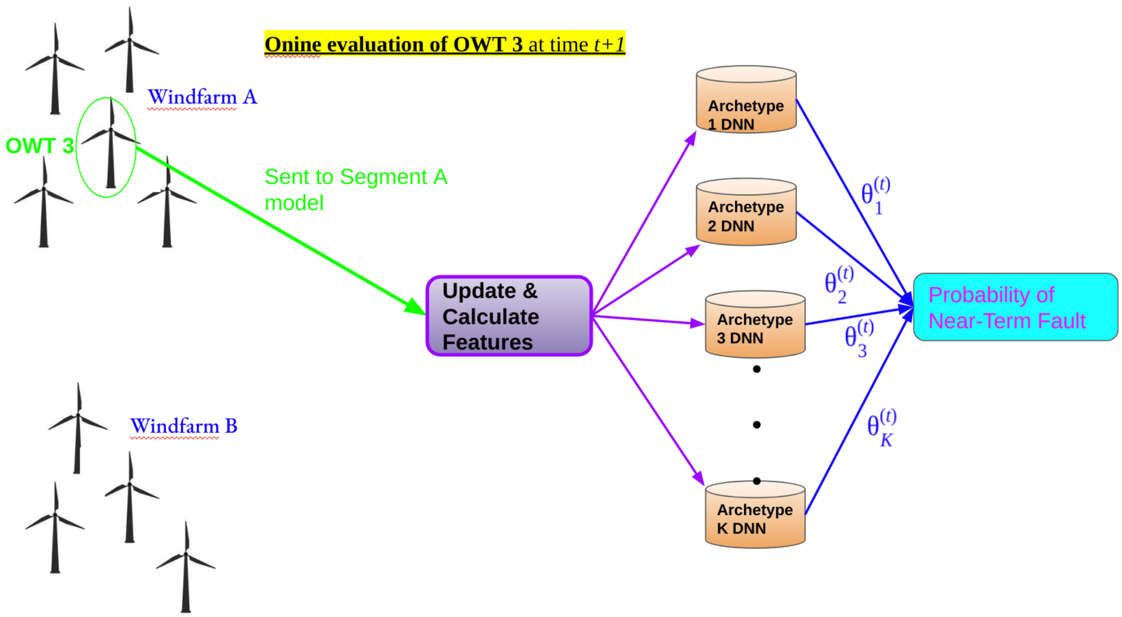

Figure 3: DNN - Ensemble

results of the developed module which allows the customer to visualize the data and then make accurate decisions. Thus, it is a vital part of the software offering.

6.1 Technical Objective 1: Task 3

This task deals with data heterogeneity issues for data clustering such that the current operating conditions of wind turbines have peer-to-peer comparisons. Our technical research and implementations in this task are beneficial to data modeling purposes (as discussed in section 5.3). A standalone module for dealing with general data heterogeneity in many different fields is a possible product of this effort. The end goal is for a user to compare a specific turbine's performance in comparison to its peers, as shown in the bar plot visualization in figure 1.

Our team of data scientists will continuously conduct reviews of data related work as the goal is to maintain the integrity of our work. Because this task is vital to our success, the PI will be managing and directing crucial technical decisions.

6.2 Technical Objective 2: Task 4

This task deals with creating modules for updating archetype models and also dealing with the sparsity of available fault data. Upon completion of this task, we expect to have a fully functional learning scheme for updating archetype models in real-time. In addition, a synthetic data generator for fault/anomalous data will result and assist in understanding what the failure space of particular wind turbines looks like. This synthetic data generator will also add value by allowing the generation of "mock" fault/anomalous data that can be used during model training. As seen in figure 1, this would allow for the identification of anomalous data vectors through the visualization engine.

The research and implementation efforts for this task, if carried beyond Phase 1, can be refined for use as standalone modules integrated in different anomaly detection settings where fault/anomalous data is sparse and data is heterogeneous. Again, the PI will be managing the process and instruct key technical decisions throughout the process.

6.3 Technical Objective 3: Task 5

This is a major task for for completing our core ML engine. The models that we will be created will be appropriately verified, tested, and fine-tuned for best performance and efficiency. We will train on the clustering and synthetic data generated explained in tasks 3 and 4 with specific goal of predicting critical turbine component failures. We will be using a combination of python the libraries: Scikit-learn, Keras, PyTorch, Numpy, and Pandas to successfully carry out the work. Our team of data scientist will closely work together to develop these tools as we have strong experience in this area.

Our results from this task would greatly benefit if carried over to Phase II, since ML models benefit from multiple iterations and increased operational data deepening the dataset. This is especially true for the sparsity in fault/anomalous data that could take years to generate. Again, the PI will manage this process and instructing critical technical decisions made along the way.

6.4 Technical Objective 4: Task 6

This task consists of organizing all components to our predictive maintenance solution in a software package that will act on real-time data from existing sensors on wind turbines and provide alerts based on the models created in task 5. We will be working closely with LEEDCo and Case Western Reserve University to receive critical feedback on our solutions and prepare changes which will increase our model quality metrics.

7 Performance Schedule

- Task 1: Milestone: Data pipeline as described in Figure 1 with a primitive ML model and visualization. Duration: July 1 2019 - July 31 2019.
- Task 2: Milestone: Library for data preprocessing required for Machine Learning. Duration: August 1 2019 - August 31 2019
- Task 3: Milestone: Create peer-peer comparison model for turbine irrespective of operating condition and weather condition. Duration: September 1 2019- October 15 2019
- Task 4: Milestone: Create models for synthetic data and verify validity. Duration: October 16 2019- November 30 2019
- Task 5: Milestone: Create ML models for anomaly detection/fault prediction December 1 2019- February 29 2020
- Task 6: Milestone:Demo the tech on live turbine data. Duration: March 1 2020 - April 15 2020

8 Facilities/Equipment

We will procure our Turbine data from LEEDCo and Case Western Reserve University. All the developmental work will be done at Bedrock Global.

9 Research Institution

N/A

10 Other Consultants and Subcontractors

Case Western Reserve University will provide the turbine sensor data including the vibration data of the structures. We have set aside $30,000 from our grant. We recently finalized the partnership and are unable to provide letter of commitment, but a letter of support is attached.

References

[1] M. Bach-Andersen, B. Rømer-Odgaard, and O. Winther. Flexible non-linear predictive models for large-scale wind turbine diagnostics. *Wind Energy*, 20.5:753–764, 2017.

[2] M. Bach-Andersen, B. Rømer-Odgaard, and O. Winther. Deep learning for automated drivetrain fault detection. *Wind Energy*, 21.1:29–41, 2018.

[3] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, 9.8:1735–1780, 1997.

[4] F. T. Liu, K. M. Ting, and Z. Zhou. Isolation forest. *Eighth IEEE International Conference on Data Mining*, 2008.

[5] M. Lukoševičius and H. Jaeger. Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, 3.3:127–149, 2009.

[6] J. Pathak and et al. Model-free prediction of large spatiotemporally chaotic systems from data: a reservoir computing approach. *Physical review letters*, 120.2, 2018.

[7] R. B. Randall. *Vibration-based condition monitoring: industrial, aerospace and automotive applications*. John Wiley & Sons, 2011.

[8] J. Tautz-Weinert and S. J. Watson. Using scada data for wind turbine condition monitoring–a review. *IET Renewable Power Generation*, 11.4:382–394, 2016.

[9] T. Wu and et al. Meta-learning autoencoders for few-shot prediction. *arXiv preprint arXiv:1807.09912*, 2018.