HOLISTIC CLASSIFICATION OF WIND TURBINE PERFORMANCE

Final Report

Author: Alec Brandel, Student, Mechanical Engineering
Company Sponsor: Invenergy, LLC
Invenergy Mentor: Jim Klus, Manager, Performance Analytics
UM Faculty Advisor: Margaret Wooldridge, Professor, Mechanical Engineering

Executive Summary

As wind energy capacity continues to grow at a rapid pace, there is a continued need to quantify and understand long-term, holistic wind turbine health and performance. Therefore, the purpose of this project is to design a system that compiles subsystem performance metrics into an overall health score. This health score should be at both the site- and turbine-level and encompass time intervals differing in size. The score should also be dynamic, and allow for the inclusion or removal of metrics as needed.

The first step in the analytical process was to benchmark existing solutions. Deviation from the power curve, mean time-to-failure and mean time-to-repair, turbine reliability, gearbox and drivetrain health, and power coefficient are all existing metrics or concepts that are commonly used to assess turbine health and performance. However, each of these metrics individually only quantify health or performance, not both, and some require analyses to be conducted on the wind turbines in addition to the data already collected. Data signals to include in the analysis were then researched and brainstormed based on existing supervisory control and data acquisition (SCADA) systems. These signals include active power, reactive power, current, voltage, blade pitch angle, tip speed ratio, and digital state.

With the selection of data signals, various metrics were calculated to quantify turbine performance such as the ratio of measured active power to rated power, the distribution of reactive power across a site, the standard deviation in blade pitch angles, the time spent in each digital state, and the tip speed ratio. These quantities plotted as a function of turbine number, time, and wind speed were then compared to one another to detect trends and discern which quantities could be used to represent others. While a correlation was found between tip speed ratio and active power ratio, it was ultimately determined that too much information would be lost in synthesizing these various metrics together. Therefore, the solution of this project is to create a health dashboard rather than a singular health score.

The proposed health dashboard contains three performance metrics: the percent time that the turbine is available to generate power but does not generate power, the active power ratio, and the standard deviation in blade pitch angle, as the pitch system was specifically highlighted by Invenergy, LLC, the company sponsoring this project, to be of interest. These three metrics would be at both the site- and turbine-level to allow for direct comparisons between sites in the wind fleet and turbines in a specific farm. The health metrics would also be presented at both the monthly and annual levels to allow for a yearly performance overview as well as an indication of the change in performance from month to month. As this dashboard is further developed and put into use, metrics can be easily added or removed to incorporate more subsystems. This dashboard therefore satisfies all of the project requirements.

Table of Contents

Problem Definition

Problem Statement

Renewable energy systems are an essential tool in decarbonizing the energy sector and combating climate change. To compete with fossil fuels and successfully meet growing demand, wind turbines need to be reliable and long-lasting. Therefore, in a continued effort to track holistic wind turbine health and performance, there is a need to design a system that compiles subsystem performance metrics into an overall turbine health score. This score should be at both the turbine-level and site-level and encompass time intervals of varying lengths. Additionally, this score should allow for the inclusion of additional metrics for future improvement.

Background

Wind turbines are a zero carbon-emitting technology that convert the kinetic energy in the wind to electrical energy. The wind spins the turbine blades, and that rotational motion turns a rotor which then turns a generator. The generator then converts the rotational motion into electrical energy which can be sent directly to the electricity grid.

Wind turbines are particularly popular for their scalability, with wind turbines varying in size from 20 W at the smallest to over 2.8 MW at the utility-scale [Small, 2021; Klus, 2021]. When compared to solar photovoltaics, which is another renewable energy source that has gained popularity, wind turbines take up far less land area for the same scale of electricity generation and are relatively more efficient [Advantages, 2021].

Wind energy capacity is rapidly growing in both the United States and around the world. In 2019, wind energy made up approximately 6.4% of global installed electricity capacity [Sönnichsen, 2019]. From 2019 to 2020, global installed wind capacity increased by approximately 122 GW, which is an increase of 19.6% [Worldwide, 2020]. In the US, wind capacity is currently 118 GW, which is a 14.2 GW, or 13.6%, increase from 2019 [U.S., 2020]. Figure 1 below shows the breakdown of US energy consumption by energy source, with wind energy encompassing 24% of renewable generation and nearly 3% of all electricity generation.

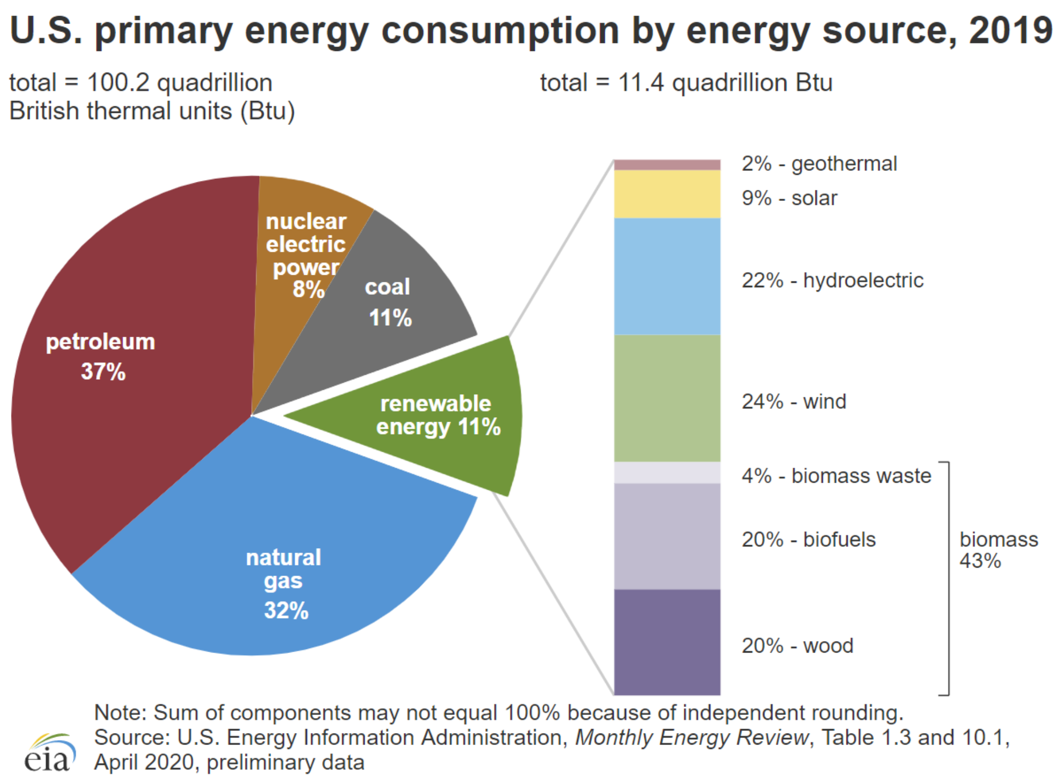

Figure 1. The above figure shows a pie chart outlining the breakdown of US primary energy consumption by energy source. The renewable energy slice is broken down into respective energy sources on the right, with each percentage representing the total percentage of that specific energy source in the renewable slice [U.S., 2020]

While wind energy serves a public need of combating climate change and producing clean energy, private companies have made large investments in the wind industry due to its increasing profitability. One such company investing in and constructing utility-scale wind farms is Invenergy, LLC, the sponsor of this project. Invenergy currently developed, owns, and operates nearly 16.7 GW of wind capacity [Invenergy, 2021].

Benchmarking

To understand current metrics for turbine health and performance, existing solutions were benchmarked. The five main metrics for turbine health documented were deviation from the power curve, mean time-to-failure and mean-time-to-repair, turbine reliability, gearbox and drivetrain health, and power coefficient.

Deviation from Power Curve

A common metric for assessing the performance of a wind turbine is the deviation from the power curve. A power curve is a plot of rated power versus wind speed. Every turbine has a power curve, but the specific power curve varies depending on the turbine type. As a wind turbine operates, the measured power generated can be plotted against measured wind speed and compared to the rated power curve. There is typically a band of acceptable power curves, and turbines that are at the edge of that band or outside of the acceptable range can be classified as underperformers [Giwhyun, 2015]. However, in studying the deviation from the power curve, it can be hard to assess changes in power curve deviation over time.

Mean Time-to-Failure and Mean Time-to-Repair

The metrics of mean time-to-failure and mean time-to-repair are closely related to one another and are common tools used in assessing turbine health. Mean time-to-failure quantifies the average amount of time it takes for a wind turbine to fail, whereas mean time-to-repair quantifies the average amount of time it takes for a wind turbine to need repair. These metrics can then be compared to the expected turbine lifetime to determine how much longer a wind turbine is expected to last. If a turbine has a lower mean time-to-failure and mean time-to-repair, then it fails more rapidly than other turbines and could be said to have lower health [Wind, 2021]. While these metrics are useful in quantifying the health of a wind turbine and extrapolating how much time is expected before the next failure or repair event, they do not indicate performance during power generation.

Turbine Reliability

Turbine reliability differs from mean time-to-failure and mean time-to-repair in that it quantifies the number of failures per year rather than the amount of time between failures [Sheng, 2015]. Turbines with a higher number of failures per year are considered to be less reliable and could be said to have lower health and performance. Similar to mean time-to-failure and mean time-to-repair, turbine reliability quantifies turbine health and does not indicate turbine performance during power generation.

Gearbox and Drivetrain Health

Considering gearbox and drivetrain health is much more difficult than the previously mentioned metrics, as it requires a vibration analysis and external studies into oil debris and gearbox temperature [Sheng, 2015]. While the previous existing metrics can all be calculated using data collected from the wind turbines and maintenance records, external companies are typically brought in to perform gearbox and drivetrain health studies [Collin, 2017]. This requires both time and money and is, therefore, less convenient than other existing methods for assessing turbine health and performance.

Power Coefficient

The power coefficient is a common metric used to assess wind turbine performance and is the ratio of the electricity produced by the wind turbine and the total energy available in the wind [Gouriérès, 1982]. While this is a useful tool in assessing the efficiency of a wind turbine, the power coefficient is also impacted by other turbine characteristics like the swept area of the blades. Therefore, the power coefficient may not be useful in drawing comparisons in performance across sites. Additionally, the magnitude of the power coefficient may not be intuitive, and may only be useful in comparing power coefficients between wind turbines.

Stakeholders

The primary stakeholders of this project are Invenergy and the wind industry as a whole. Invenergy is the sponsor of this project and has provided both mentorship and data for analysis. Additionally, the results of this project will be used specifically by Invenergy after completion and the system created was made to integrate into Invenergy's current operational system.

This project is also important to the overall wind industry, as working to further the understanding of wind turbine health and performance would benefit the entire industry, especially if the results could be used to predict potential future issues.

Discussion

Methods

The first step in the project, aside from researching existing solutions, was to determine which data signals should be included in the analysis. To get an idea for all potential signals, research was done into common supervisory control and data acquisition (SCADA) system monitoring variables. A flowchart of SCADA monitoring variables within their corresponding turbine subsystems is shown in Figure 2 below.

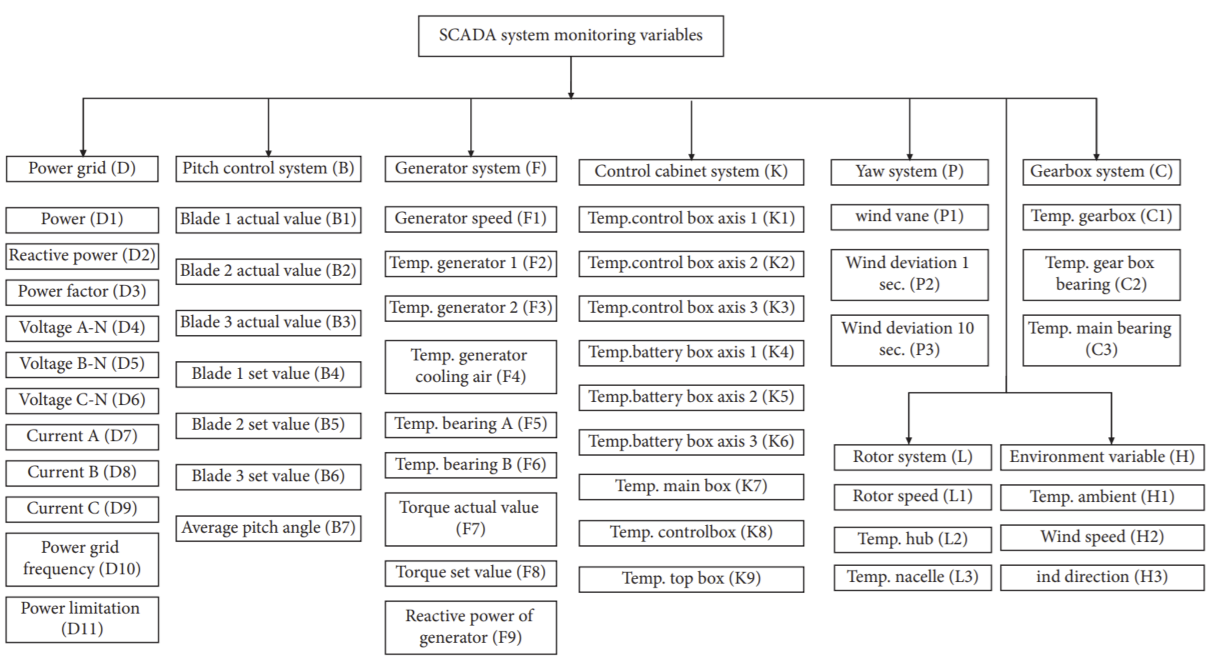

Figure 2. The above figure shows a flow chart of the data signals traditionally monitored by a SCADA system, with

the top element in each section being the title of the subsystem encompassing the respective set of signals [Tao, 2019].

Each of the SCADA data signals were then discussed with Invenergy to determine which signals are most important to initially include in the analysis. A priority level for each signal was then generated based on importance to Invenergy [Klus, 2021]. Data signals included in the analysis and their respective priority levels are shown in Table 1 below.

Table 1. The following table contains each of the data signals to be included in the subsequent project analysis with their corresponding subsystem and priority level.

Subsystem	Data Signal	Priority Level
Power Grid	Active power	Highest
	Reactive Power	Medium-Low
	Current	Medium
	Voltage	Medium
Pitch Control	Blade pitch angle	Medium-High
	Tip Speed Ratio	High
Digital Signals	Digital state	High

In addition to selecting data sources for inclusion in the project, the scale of the data also needed to be considered. Data is collected at both instantaneous intervals and averaged across 10-minute intervals [Klus, 2021]. To get an accurate understanding of holistic health and performance, data from one or more years needs to be considered. With data commonly stored as 10-minute averages, one year of data for one turbine consists of over 50,000 data points. Multiplying this by the number of data signals and the number of turbines at a site yields over 22 million data points to consider for one year of analysis. Not only can this be overwhelming to parse through and study, but this scale of data can have slow computation times in most computational programs like Python and MATLAB. Therefore, part of the analysis needs to involve limiting the granularity of the data, as considering each 10-minute average is too fine for a holistic analysis. This can be done by considering various common turbine performance metrics and averaging data over larger, monthly time intervals. Another important consideration is the potential for the collected data to change format over time. Since the SCADA system is widely used across the wind industry, it is very unlikely that the scale at which data is collected or the signals measured will change in the future.

With the different data signals selected, various sub-analyses were conducted in an effort to quantify performance and digest the data. These include calculating the ratio of measured active power to rated power, the distribution of reactive power across a site, the standard deviation in blade pitch angles, the time spent in each digital state, and the tip speed ratio. Data was provided by Invenergy for one of their wind farms. However, for the sake of confidentiality, the name and location of the site remained anonymous. Additionally, no data or quantifiable results will be shown in the following discussion of analyses and results. All data signals had measurements provided as 10-minute averages across one year.

Each of the subsequent analyses were performed in Jupyter Notebooks, which uses Python as the scripting language. Python was chosen for its ability to easily parse through large sets of data, and Jupyter Notebooks was selected as a platform due to its ability to run portions of code multiple times without running the entire script.

Data Filter

The first step in the analytical process was to filter out known offline periods. While the duration of offline periods is an area of interest, active power ratio, reactive power distribution, standard deviation in blade pitch angle, and tip speed ratio should be calculated by only considering periods when the turbine is generating power. This is due to the fact that the main focus of these metrics is to assess the performance of the wind turbine as it is generating power. Instead of filtering the data before calculating each of these metrics, a conditional statement was added to each individual calculation. For example, the active power ratio was only added to the monthly average if the digital state at that specific turbine and timestamp indicated that the turbine was generating electricity. Therefore, timestamps where a turbine was not generating power were not incorporated into monthly and annual averages. An example of what this conditional statement looks like is shown in Figure 3 below, where generating_state is a variable that represents the digital state at which the turbine is generating power. In implementation, the phrase "perform calculation" would be replaced with the calculation that needs to be conducted, like computing the active power ratio.

```
if filtered_state.iat[x,y] == generating_state:
    perform calculation
```

Figure 3. The above figure shows a representation of the conditional statement applied to each calculation in the Python script, which acts as a data filter.

Active Power Ratio

Active power ratio, or the ratio of measured active power to rated active power, was the first metric to be calculated. Essentially, the active power ratio shows the fraction of rated power that a turbine is operating at, with turbines theoretically expected to generate as close to rated power as possible. For turbines that are performing well, one would expect the active power ratio to be approximately equal to 1.

To calculate the active power ratio, the rated power first had to be determined from the power curve. A power curve is a plot of rated active power versus wind speed and shows the rated power that a turbine is expected to operate at for given wind speeds. Each turbine has a power curve provided by the manufacturer, but the power curve differs based on turbine size and type.

Using the wind speed measured by the anemometers on each wind turbine at the site, the rated power was calculated for each timestamp by taking the wind speed and linearly interpolating on the power curve to find the corresponding rated power. The active power measured by each turbine at each timestamp was then divided by the calculated rated power at that same timestamp to generate the active power ratio.

Monthly and annual averages were also calculated for each turbine by summing the measured active power across each month and summing the rated active power across each month. These sums were then divided by the total number of timestamps in the month when the turbine was generating power to yield averages. These averages were then divided by one another to calculate the ratio of average measured active power to average rated active power. This same process was used to calculate the average active power ratio across the year. To generate the site level active power ratio at the monthly and annual level, the respective monthly and annual active power ratios for each turbine were averaged together.

Lastly, a study was done into identifying instances of subtle underperformance, where a subtle underperformance is a case in which a wind turbine is generating between 80-90% of rated power. The total instances of subtle underperformance per turbine were calculated by iterating through each timestamp and recording whether or not the active power ratio was between 0.8 and 0.9. This area is of ongoing study, and results cannot be shared due to confidentiality concerns. Additionally, the bounds that determine whether or not a turbine is underperforming may be adjusted depending on results from other sites.

Reactive Power Distribution

The reactive power distribution was considered in an exploratory manner rather than in an effort to quantify turbine performance. Reactive power deals with the distribution of active power and managing voltages across the grid. In the case of a wind farm, reactive power assists in setting the voltage at which active power is transmitted to the grid and ensures that the electricity generated is transmitted efficiently [Chen, 2012]. Intuitively, it makes sense for the reactive power of a site to be distributed among the wind turbines evenly, so that no individual turbine experiences a drastically higher or lower reactive power. Disproportionate reactive power across a site could indicate issues at either the turbine or the site level related to power transmission.

To quantify the reactive power distribution across a site, a concept similar to the active power ratio was used. However, reactive power differs in that the reactive power should be approximately equal across all wind turbines in a site at each given timestamp. Additionally, there is not a single value for rated reactive power and there does not exist a reactive power curve since the reactive power depends on grid transmission, not wind speed. Therefore, the "rated value" was substituted for the average reactive power across the site for a respective timestamp.

Once the average reactive power was calculated at each timestamp, each measured reactive power was divided by the average at each timestamp to generate a ratio of measured reactive power to the average across the site, or the reactive power ratio. In an ideal scenario, the reactive power ratio would equal 1 at each timestamp, indicating that the reactive power is distributed evenly across the site.

Monthly and annual averages were calculated using the same approach as the monthly and annual active power ratios. The average reactive powers of the site at each timestamp were averaged together across the month and year. The reactive power of each turbine was also averaged across each month and the entire year and then divided by the site average to generate the monthly and annual reactive power ratios.

Standard Deviation in Blade Pitch Angle

Each utility-scale wind turbine contains a pitch system. This system pitches the blades of the wind turbine depending on the wind speed to maximize the energy extracted from the wind. Therefore, the pitch system plays a key role in how efficiently the wind turbine generates energy and is indicative of wind turbine performance.

For most pitch systems, the blades are meant to pitch to the same degree. If the three blades do not pitch correctly, this could lead to diminishing efficiency. To determine whether or not the blades are pitching identically, the standard deviation in the pitch angles of the three blades for each turbine at each timestamp was calculated. Turbines with the highest standard deviation in blade pitch angles may be most likely to have efficiency losses caused by the pitch system.

Monthly and annual averages were also calculated for each turbine by simply summing the standard deviation in blade pitch angles and dividing by the total number of timestamps where the turbine is generating power across each month and across the year.

Digital States

Wind turbines typically have a set of digital states that describe the operating state of the turbine. For example, the turbine would read out a certain value if it is operating normally and another if it is under repair. While these digital states vary depending on turbine type and manufacturer, the existence of some form of digital state is consistent among all turbines.

In looking at digital state data for one site taken as 10-minute averages across one year, the number of offline periods and the duration of those offline periods were studied on both a monthly and annual basis. This provides information regarding the percent time each turbine was generating power, offline, or available but not generating. The case of the turbine being available to generate power but not generating was of particular interest, as those represent periods in which the turbine could be generating power, but is not doing so. To calculate the fraction of time in each digital state, all timestamps for each turbine were looped through and a running sum was kept of the number of timestamps each turbine was in each state. The sum of timestamps in each state was then divided by the total number of timestamps that constitute one year to determine the percentage of time that each turbine spends in each digital state. A similar process was used to calculate the amount of time each turbine spends in each state on a monthly basis. Turbines that spend a higher percentage of time not generating power would then be flagged as underperforming.

Tip Speed Ratio

The tip speed ratio is the ratio of the tangential speed of the tips of the wind turbine blades to the wind speed. The tip speed ratio is commonly used to quantify wind turbine efficiency, and different turbines have different optimal tip speed ratios. In the case of this analysis, the tip speed ratio of each turbine at each timestamp was calculated by dividing the product of rotor speed and blade length by the measured wind speed. This calculation was conducted at each timestamp for each turbine, as well as averaged across the month and year using the same process as used for active power and reactive power.

Health Score Synthesis

Once the various calculations were performed, ways to synthesize the metrics into a holistic health score were brainstormed. A key consideration in synthesizing scores together was how much information could be lost in the process. With that consideration in mind, plots of active power ratio, standard deviation in blade pitch angles, time spent in each digital state, and the tip speed ratio as functions of time, wind speed, and turbine number were compared to one another in an attempt to discern trends and determine whether or not certain metrics could be used to represent others.

Results

Upon comparing plots of active power ratio, reactive power distribution, standard deviation in blade pitch angles, time spend in each digital state, and tip speed ratio as a function of time, wind speed, and turbine number, it was quickly evident that there was not one metric that could be used to represent trends in the others. The one exception to this finding was that there appeared to be a correlation between tip speed ratio and active power ratio. This correlation is shown in Figure 4 below, where the bar chart showing the annual average tip speed ratio for each turbine is ranked in order from highest active power ratio to lowest active power ratio. It is important to note that the following plots have had their y-axis quantities removed so as to protect the confidentiality of the data and results.

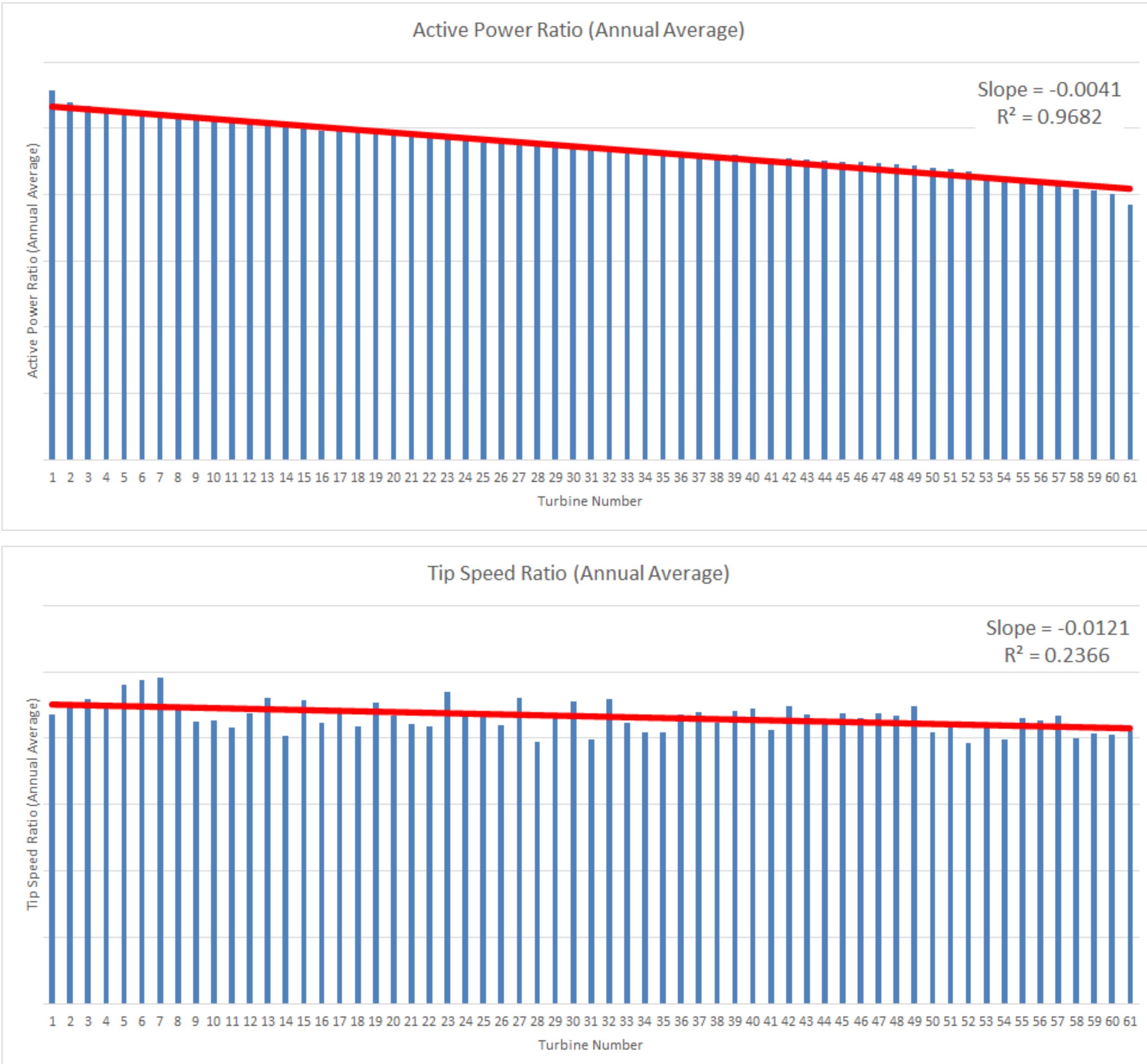

Figure 4. The above figure shows two plots, the first being a bar chart of the annual average active power ratio for each turbine ranked in descending order (top). The second plot is a bar chart of the annual average tip speed for each turbine ranked in the same order as the active power ratio plot (bottom). The slope and R^2 value for a linear trendline is shown on each respective plot as well. The y-axis quantities have been removed from each plot and the turbines have been renumbered to protect the confidentiality of the data and results.

As evident in the above figure, turbines with a lower active power ratio also tend to have a lower tip speed ratio. The trendlines on each bar chart have a negative slope, confirming this correlation. Further confirming this correlation is the fact that the R^2 values, which are a measure of how close the trendline fits the data, are within an order of magnitude of 1, where an R^2 of 1 indicates a perfect fit. This correlation makes sense intuitively, as tip speed ratio is a measure of the efficiency of the wind turbine, so turbines with lower efficiency should theoretically yield lower power output than expected.

While tip speed ratio and active power ratio follow a similar trend, neither metric appears to provide an understanding of the amount of time spent in each digital state, the standard deviation in blade pitch angles, or the reactive power distribution. Figure 5 below shows the standard deviation in blade pitch angles, the fraction of time each turbine is available to generate power but is not generating power, and the reactive power ratio for each turbine averaged across the year. Again, the turbines are ranked in order from highest active power ratio to lowest active power ratio, so a trendline with a negative slope and an R^2 value within an order of magnitude of 1 would indicate a correlation between the respective metric and active power ratio.

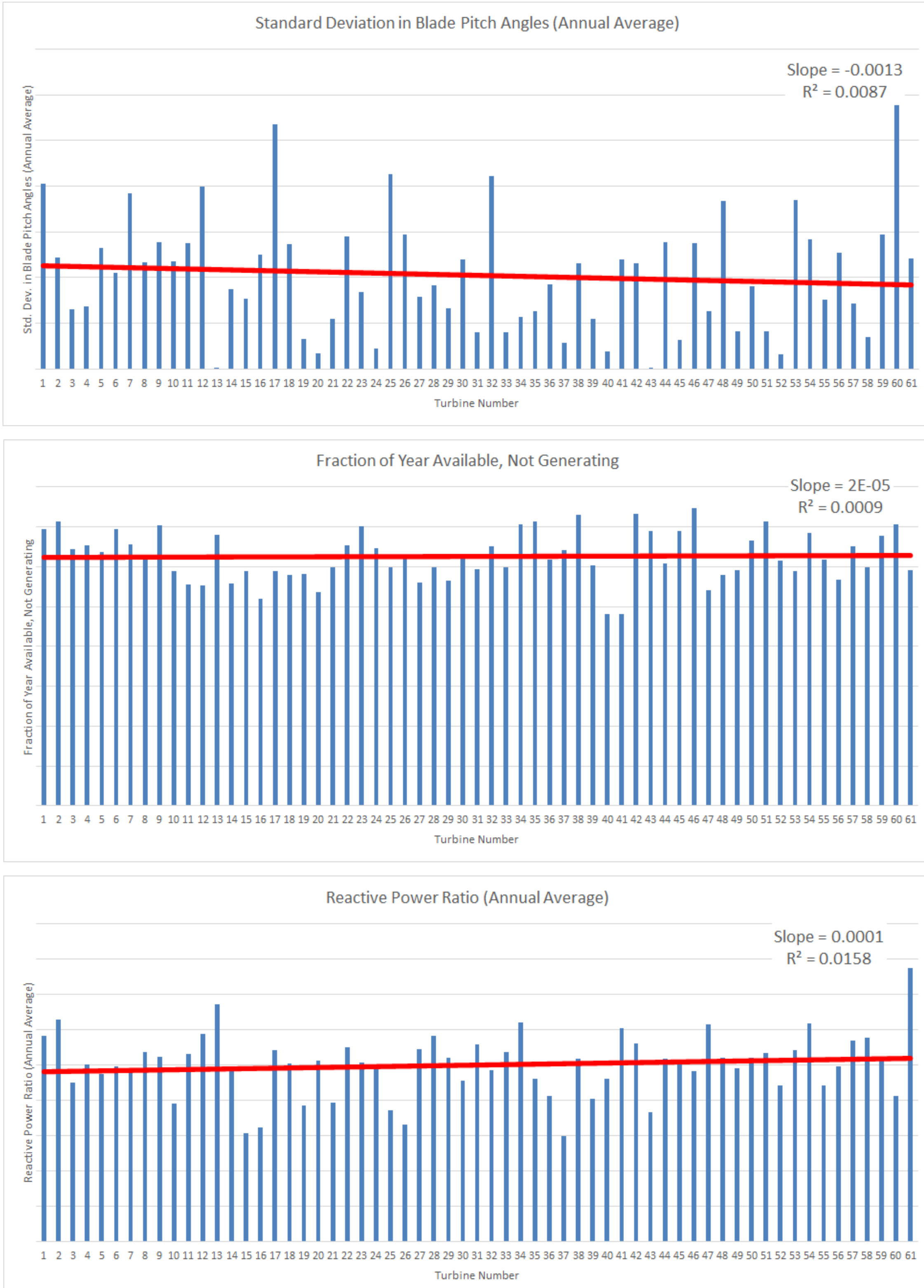

Figure 5. The above figure shows bar charts of the standard deviation in blade pitch angles (top), the fraction of year available, but not generating (middle), and the reactive power ratio (bottom) averaged across the year for each turbine. A linear trendline is shown on each plot with the corresponding slope and R^2 value. Y-axis values have been removed from each bar chart for data confidentiality purposes.

As evident in the above figure, there does not appear to be a correlation between either of the three metrics and the active power ratio. Both the fraction of the year the turbine is available but not generating and the reactive power ratio appear to have a trendline with a positive slope and very small R^2 value, indicating a potentially inverse relationship. However, with such a small R^2 value of multiple orders of magnitude less than 1, it is impossible to discern any relationship with certainty. Additionally, while the standard deviation in blade pitch angles shows a negative trend, the R^2 value is nearly two orders of magnitude less than 1, implying that again a correlation cannot be made with certainty. Therefore, the active power ratio does not provide any understanding as to the amount of downtime a turbine has, the standard deviation in the blade pitch angles, or the reactive power distribution across a site.

The lack of correlation between the fraction of time available but not generating and active power ratio makes sense intuitively, as the active power ratio is calculated only for times when the turbine is generating power, so it would provide no indication of how long the turbine is generating power. The lack of correlation between standard deviation in blade pitch angles and active power ratio also makes sense intuitively, but the intuition is not as obvious. Although the plot of standard deviation in blade pitch angle in Figure 5 shows a larger difference between the standard deviation of one turbine to another, the removal of the y-axis also removes any indication as to what the scaling is on the plot. The standard deviations for each turbine are all within 1.2 degrees of one another, which is a very small range considering turbine blades can pitch greater than 100 degrees. With relatively small deviations in blade pitch angles on average, it makes sense that this would not have a significant impact on the active power ratio. However, it is still unclear whether subtle variances in blade pitch angle would lead to longer-term issues, especially in the pitch system. Lastly, the lack of correlation between reactive power ratio and active power ratio also makes sense intuitively, as the reactive power needs of each turbine depend on characteristics of the energy grid and less on the performance of the turbine. However, more research and analysis will need to be done on reactive power to certify this claim.

Since the amount of time a turbine is able to generate power but does not do so is evidently independent of the active power ratio and the blade pitch system was marked as a specific area of interest by the sponsor, it is clear that synthesizing these metrics into one health score would lead to lost information. Therefore, instead of generating a turbine health score, the solution to this project is to establish a turbine health dashboard that shows the fraction of time a turbine is available but not generating, the active power ratio, and the standard deviation in blade pitch angle. These three metrics would be shown at both the turbine- and site-level and span both monthly and annual time periods. A flowchart depicting the spatial and temporal relationship within this dashboard is shown in Figure 6 below.

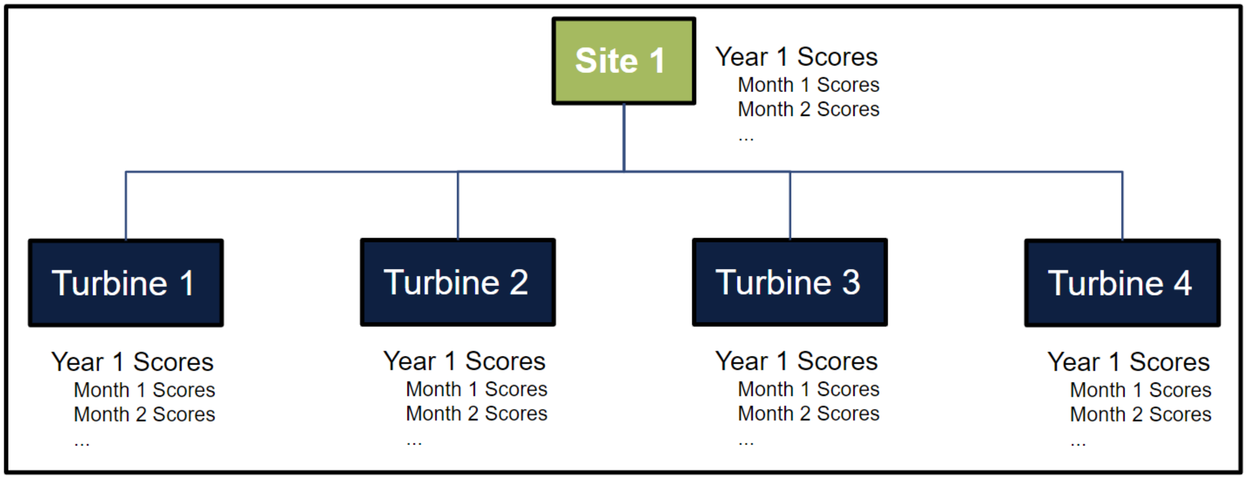

Figure 6. The above figure shows a flowchart depicting the spatial and temporal hierarchy associated with the proposed health dashboard. Scores are generated at both the site- and turbine-level as well as the monthly and annual level.

While the creation of the dashboard is a future step in the project, the elements to be included in the initial iteration of the dashboard have been determined. Along with displaying values for the three metrics on a monthly and annual level, there would also be an option within the dashboard to view plots showing comparisons between all turbines in a site for each of the three metrics and the change over time of each metric at the turbine-level. The format of these plots would be similar to the ones in Figure 4 and Figure 5 above. As the dashboard is created and used, it can be modified to add or remove metrics and plots at any point in the future.

Recommendations and Next Steps

In continuing this project, the first next step would be to modify the metric displaying the fraction of time each turbine is available to generate power but does not generate power with how much energy could have been generated if the turbine were generating power. This would be done using a similar method as used to calculate the rated power at each timestamp. Using wind speed measurements, the rated power can be interpolated from the power curve and then summed across all timestamps in which the turbine is not generating power. This would then give the total power that could have been generated each month and across the year to better quantify the losses experienced. Taking this calculation one step further, the total power could be multiplied by the average revenue associated with one unit of power at each site to quantify how much revenue is lost by not generating power.

Another next step to expand the results would be to include data from more wind farms across multiple years into the analysis. Sampling from a larger pool of data would allow for broader comparisons to be made and maximize the chances of detecting trends in data. Additionally, including data from more sites would allow for comparisons between sites to be studied. Current and voltage, which were previously identified as signals that could be included in the analysis, should also be considered for further development of the health dashboard.

Although the content included in the dashboard has been determined and there is a system to calculate the necessary health metrics, this system needs to be modified to integrate into Invenergy's current operational system. An interface for the dashboard also needs to be created and tested with potential end-users to maximize usability. The dashboard has also been made to intentionally be modified. As the results of the dashboard are studied further by potential end-users, metrics may be added or removed from it depending on their necessity.

To further study long-term trends, the number of instances of high or low reactive power and high standard deviation in blade pitch angles across each month and year should be studied. While this has already been considered briefly, more research needs to be done to determine the threshold that defines what high reactive power, low reactive power, and high standard deviation in blade pitch angles mean. With further analysis of potential trends, a study should be done into whether or not the metrics in the health dashboard can be used to pre-emptively detect and solve issues.

Conclusion

Ultimately, the proposed health dashboard successfully meets the project requirements and objectives. In displaying the essential performance metrics of the fraction of the year that the turbine is available but not generating power, the active power ratio, and the standard deviation in blade pitch angles at both the turbine- and site-level, the dashboard provides a holistic view of turbine health and performance. Additionally, the incorporation of multiple metrics allows for additional metrics to be included or existing metrics to be removed as the health dashboard is used in industry. Lastly, the dashboard displays each metric at both the monthly and annual levels, allowing for a broad view as well as an understanding of changes in performance from month to month.

Upon looking at the dashboard, an end-user would be able to immediately see the fraction of the year that a turbine was able to generate power, but was not generating power. Turbines that have a higher fraction are generating power less on average than neighboring turbines and may be taking longer to ramp up to production or experiencing more curtailment. The active power ratio then shows how close to rated power a turbine, on average, is producing when it is generating power. This is a direct representation of the average efficiency of a wind turbine, and turbines with lower active power ratios are performing worse, on average, than neighboring turbines. Lastly, the standard deviation in blade pitch angle gives a direct view of the performance of the pitch system, which was identified to be a top priority by Invenergy, LLC, the sponsor of this project.

While specific results of the analyses performed in this project cannot be shared for data sensitivity reasons, the process used to analyze the data is part of the outcome. Additionally, trends observed in the process of synthesizing various performance metrics into an overall health score provided insight on their own. For example, the correlation between active power ratio and tip speed ratio reinforces the intuition that using different methods to measure wind turbine efficiency should yield similar results.

As this project is continued, the fraction of year available but not generating metric will be modified to better quantify how much energy is lost due to the lack of production. Additional subsystems, such as the yaw system, will also be studied for potential inclusion in the health dashboard as the scope of the project increases both spatially and temporally to encompass multiple wind farms and multiple years of data.

Acknowledgments

This project would not have been possible without sponsorship from Invenergy, LLC, and support from my mentor at Invenergy, Jim Klus. Jim met with me consistently throughout the project, providing me with data and feedback as to whether or not the project was in line with Invenergy's requirements.

I would also like to thank the faculty advisor for this project, Professor Margaret Wooldridge. Professor Wooldridge provided continuous support and advice, especially regarding technical communication and managing data-intensive projects.

References

"Advantages and Challenges of Wind Energy." *Energy.gov*, Office of Energy Efficiency and Renewable Energy, www.energy.gov/eere/wind/advantages-and-challenges-wind-energy.

Bilir, Levent, Mehmet İmir, Yılser Devrim, and Ayhan Albostan. "An Investigation on Wind Energy Potential and Small Scale Wind Turbine Performance at İncek Region – Ankara, Turkey." *Energy Conversion and Management* 103 (October 1, 2015): 910–23. https://doi.org/10.1016/j.enconman.2015.07.017.

Byon, Eunshin, et al. "Adaptive Learning in Time-Variant Processes With Application to Wind Power Systems." *IEEE Transactions on Automation Science and Engineering*, vol. 13, Apr. 2016.

Chen, Xu, et al. 2012, *Reactive Power Compensation and Energy Storage in Wind Power Plant*.

Collin. “Component Health Monitoring for Wind Turbines.” *LORD Sensing Systems | Innovating Together*, 26 May 2017, 16:54, www.microstrain.com/blog/edit-blog-component-health-monitoring-wind-turbines.

Dao, C, Kazemtabrizi, B, Crabtree, C. Wind turbine reliability data review and impacts on levelised cost of energy. *Wind Energy*. 2019; 22: 1848– 1871. https://doi.org/10.1002/we.2404

Giwhyun Lee, Yu Ding, Marc G. Genton & Le Xie (2015) Power Curve Estimation With Multivariate Environmental Factors for Inland and Offshore Wind Farms, Journal of the American Statistical Association, 110:509, 56-67, DOI: 10.1080/01621459.2014.977385

Gouriérès Le Désiré. *Wind Power Plants: Theory and Design*. Pergamon Press, 1982.

Hamilton, Sofia, et al. “How Does Wind Project Performance Change with Age in the United States?” Lawrence Berkeley National Laboratory. 5 Apr. 2020.

“Invenergy: Wind.” *Invenergy Main*, Invenergy, invenergy.com/what-we-do/wind.

Klus, Jim. Weekly Meeting, 2021.

Lee, Giwhyun, et al. Wiley Online Library, 2014, *A Kernel plus Method for Quantifying Wind Turbine Performance Upgrades*.

Mishnaevsky, L, Thomsen, K. Costs of repair of wind turbine blades: Influence of technology aspects. *Wind Energy*. 2020; 23: 2247– 2255. https://doi.org/10.1002/we.2552

Parsons, Taylor. “SCADA-Based Energy Assessments.” DNV-GL. 30 Sept. 2020.

Sheng, Shuangwen. “Prognostics and Health Management of Wind Turbines: Current Status and Future Opportunities.” National Renewable Energy Laboratory. National Renewable Energy Laboratory, 14 Dec. 2015, Sao Paolo, Brazil.

“Small Wind Guidebook.” *WINDExchange*, 2021, windexchange.energy.gov/small-wind-guidebook.

Sönnichsen, Published by N., and Feb 11. “Global Electricity Capacity by Source 2019.” *Statista*, 11 Feb. 2021, www.statista.com/statistics/267358/world-installed-power-capacity/.

Tao, Liang, Qian Siqi, Yingjuan Zhang, and Huan Shi. “Abnormal Detection of Wind Turbine Based on SCADA Data Mining.” Edited by Michael Z. Q. Chen. *Mathematical Problems in Engineering* 2019 (August 7, 2019): 5976843. https://doi.org/10.1155/2019/5976843.

"U.S. Energy Information Administration - EIA - Independent Statistics and Analysis." *U.S. Energy Facts Explained - Consumption and Production - U.S. Energy Information Administration (EIA)*, U.S. Energy Information Administration, Apr. 2020, www.eia.gov/energyexplained/us-energy-facts/.

Wind Energy Benchmarking Services, *Understanding Out of Warranty Reliability*.

Wiser, Ryan, et al. "Wind Energy Technology Data Update: 2020 Edition." Lawrence Berkeley National Laboratory. 20 Apr. 2021.

"Worldwide Wind Capacity Reaches 744 Gigawatts – An Unprecedented 93 Gigawatts Added in 2020." *World Wind Energy Association*, 24 Mar. 2021, wwindea.org/worldwide-wind-capacity-reaches-744-gigawatts/.

Wright, Sally. DNV GL, 2017, *Definitions of Availability Terms for the Wind Industry*.