

PAPER • OPEN ACCESS

On the Sensitivity of Wind Turbine Failure Rate Estimates to Failure Definitions

To cite this article: F Anderson *et al* 2023 *J. Phys.: Conf. Ser.* **2626** 012025

View the [article online](#) for updates and enhancements.

You may also like

- [Disruption of artificial lipid bilayers in the presence of transition metal oxide and rare earth metal oxide nanoparticles](#)
Shiv Acharya, Bin Lu, Shayson Edwards et al.
- [REDSHIFT MEASUREMENT AND SPECTRAL CLASSIFICATION FOR eBOSS GALAXIES WITH THE REDMONSTER SOFTWARE](#)
Timothy A. Hutchinson, Adam S. Bolton, Kyle S. Dawson et al.
- [Target Selection and Validation of DESI Luminous Red Galaxies](#)
Rongpu Zhou, Biprateep Dey, Jeffrey A. Newman et al.



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

On the Sensitivity of Wind Turbine Failure Rate Estimates to Failure Definitions

F Anderson¹, R Dawid, D McMillan², D García Cava¹

¹ University of Edinburgh, King's Buildings, West Mains Road, Edinburgh, Scotland EH9 3JG

² University of Strathclyde, 16 Richmond Street, Glasgow, Scotland G1 1XQ

E-mail: F.J.Anderson@sms.ed.ac.uk

Abstract. This study presents a wind turbine reliability analysis at the turbine and assembly level. It is concerned with the uncertainties associated with data-processing for wind turbine failure rate figures. These uncertainties are prominent in discussions of failure data in the literature. In particular, the influence of different failure definitions on failure rate estimates are investigated. The baseline estimate is 9.06 failures per turbine per year. This figure changes significantly when introducing a lower downtime limit, repair limit or limit on time between subsequent downtimes of the same turbine for a downtime event to be considered a failure. It changes significantly depending on which maintenance actions are categorised as corrective and by what data points represent an intervention. From the one dataset analysed here, results show derived failure rates ranging from below 1 failures per turbine per year to over 10 failures per turbine per year using failure definitions which have previously been used in the literature. When restricting failures to those that can be attributed to a particular assembly, the failure rate estimate reduces to 7.47 failures per turbine per year. Assemblies that fail the most are the frequency converter (at around 1 failures per turbine per year) and the cooling system (at around 0.9 failure per turbine per year). The gearbox, blades, yaw system and hydraulic group were the next most frequently failing assemblies.

1. Introduction

The offshore wind industry has been growing rapidly over the past couple of decades. Its growth is expected to continue accelerating in the the next couple of decades. In order for this trajectory to continue, the trajectory of decreasing costs for offshore wind likewise must continue. A key cost driver is operations and maintenance (O&M). Effective O&M strategy is based on preventing or reacting quickly to turbine failures or instances of non-operability. To make this process more cost effective requires a deeper understanding of wind turbine failures. It is very valuable to all stakeholders in an offshore wind farm to know how often a wind turbine fails and which assemblies are faulty when it does. Assembly failure rates can be obtained from the analysis of historical failure data. This either comes in the form of maintenance logs of technician repairs or alarm codes from turbine SCADA systems; or some combination of the two. However, the research space of wind turbine reliability analysis is inconsistent, due to inconsistent data treatment and failure definitions. This inconsistency manifests itself as a considerable uncertainty in published failure data. The purpose of this study is to investigate the uncertainty imposed on wind turbine reliability analysis through inconsistent data treatment.

The rest of the paper is structured as follows. Section 2 presents a problem statement, containing brief reviews on wind turbine reliability analysis; inconsistencies of data treatment



and standards for data collection. Section 3 details the data mining methodology, containing a description of the dataset used in this analysis, a methodology for identifying failures and a definition of technical availability. Section 4 details sources of uncertainty for the failure definition we use in the analysis. Section 5 presents results; section 6 a discussion of the results and their implication and section 7 a general conclusion.

2. Problem Statement

2.1. Wind Turbine Reliability Analysis

Wind turbine reliability analyses are an integral part of learning from previous experience in the industry. They play a crucial role in reducing costs both at the design and operations phase of a wind farm. They indicate to Original Equipment Manufacturers (OEMs) where resources should be directed to improve reliability, and indicate to wind farm operators where to direct their resources and how to optimise their operational strategy. The value of reliability analyses have also been heightened by a prevailing lack of transparency in the industry. Strict confidentiality practices are put in place by OEMs which they consider to protect their interests, but hampers every other stakeholder in making effective decisions to cut costs.

Despite this hindrance, there are several reliability analyses published in the literature. These are comprehensively reviewed by Pfaffel et al. [1] and Cevasco et al [2]. Since these review studies provide a wealth of information on the subject, we will not conduct a detailed review here. However, two points are worth highlighting in the context of this analysis:

- 2.1 (i) There are only two studies listed by Pfaffel et al. which provide failure rates for offshore turbines [3, 4]. Reliability analyses for offshore wind turbines are therefore especially valuable.
- 2.1 (ii) Despite the efforts of those review studies to extract trends from the publicly available data, the published failure rate figures are difficult to compare because the research space is so inconsistent. This inconsistency stems not only from different wind turbine concepts, environmental conditions, wind turbine ages and installation dates, but also from inconsistent data treatment. Inconsistent data treatment is the focus of this study.

2.2. Inconsistency of Data Treatment

Leahy et al. [5], Reder et al. [6] and Hahn et al. [7] all highlight the numerous issues associated with treatment of wind turbine reliability data. These can be summarised as follows:

- 2.2 (i) General lack of failure data. See section 2.1.
- 2.2 (ii) The lack of a standard failure definition in the wind industry [5]. What a researcher means when they say a turbine 'fails' varies from one study to the next. Often, researchers are restricted by the format and detail of their available dataset.
- 2.2 (iii) The lack of a standard taxonomy in the wind industry [6, 7]. There is also a large variation in how published failure rates are categorised by assembly.
- 2.2 (iv) Disparate data sources [5, 7]. Points (ii) and (iii) both stem from and influence a lack of consistency in exactly what reliability analysis looks like, to the point where relevant stakeholders (particularly operators) do not know the value of their own data.
- 2.2 (v) General data quality issues [5, 7]. Failure data also tends to be incomplete and significant manual effort is required to draw valuable conclusions.

All of these points produce a measure of uncertainty in wind turbine failure statistics. All of them taken together produce a research space in there is significant uncertainty associated with failure rate figures, and this uncertainty is rarely addressed. However, the issue of uncertainty has begun to draw attention in recent years. In particular, Seyr & Muskulus, Scheu et al.,

Dao et al. and Martin et al. all provide methodologies for exploring uncertainty in reliability data [8–11].

This study addresses point (2.2.2) of the above list - namely, the lack of standard failure definition in the industry. It does so by performing a sensitivity analysis of failure rate estimates based on different failure definitions using real-world data.

2.3. Data Collection Standards

Hahn et al. [7] and Leahy et al. [5] both provide reviews of standards for data collection, processing and categorisation for wind turbine reliability from 2017 and 2019, respectively. Main considerations are given to the following two factors:

- 2.3 (i) **Taxonomy definition.** The Reliawind Taxonomy [12] is popular among academics: it is both specific enough to provide insight and simple to apply. The Reference Designation System for Power Plants (RDS-PP) is more popular in industry, especially in recent years [13]. The North American Electric Reliability Corporation provides the Generating Availability Data System (NERC-GADS) [14] which also sees use in industry.
- 2.3 (ii) **Data Collection Guidelines.** IEA Task 33 recommends the development of a wind-centric guideline based on ISO 14224 [15] and ZEUS [16] standards for failure/fault, maintenance and inspection data. These both provide detailed guidelines, but are not wind-specific. In contrast, IEC 61400 [17] provides standards for operational data (describing the operational state of the turbine) that are widely accepted by industry.

As is common among many operational datasets, the dataset used in this analysis (see section 3.1) does not conform to the above recommended practices for data collection.

3. Data Mining Methodology

3.1. Dataset Description

This analysis is based upon the same dataset as described previously by Anderson et al. [18,19]. The dataset contains operational data provided by a large offshore wind farm, consisting of a fleet of geared wind turbines with a multi-MW power rating. It consists of approximately 600 turbine years of data. It should be noted that all wind turbines are from the same manufacturer. Failure rates (especially at the assembly-level) are likely to show significant variations between different manufacturers.

We provide a detailed metadata table for the dataset in reference [19]. Those data-tables which are relevant for a reliability analysis are summarised in table 1, which presents the information derived and disadvantages of each data-table. Work procedures is a dataset which allows more thorough investigation of turbine failures, and was not incorporated into our previous research. Together with tasks/task types, it allows the failure rate of assemblies to be estimated.

3.2. Identification of Turbine-Level Failures

Turbine-level failures are identified via the process laid out in reference [18]. Namely, a *Downtime Catalogue* is constructed by cross referencing the data tables *Tasks*, *SCADA and Operations Planned Movements*. In the creation of the *Downtime catalogue*, three things are accomplished:

- 3.2 (i) Turbine downtime events are extracted from *SCADA*.
- 3.2 (ii) Those downtime events which are accompanied by an intervention from the maintenance team are extracted from *SCADA and Operations Planned Movements*.
- 3.2 (iii) Corrective maintenance interventions are extracted from *Task Types* together with *SCADA and Operations Planned Movements*.

Table 1: Summary of the data-tables relevant to a reliability analysis.

O&M Data Type	Information Derived	Disadvantages
A. Work Procedures	<ul style="list-style-type: none"> • Text description of work carried out 	<ul style="list-style-type: none"> • Sometimes difficult to map to assembly taxonomy
B. Tasks/Task Types	<ul style="list-style-type: none"> • Task descriptions • Task categories 	<ul style="list-style-type: none"> • Most corrective task descriptions only contain an alarm code • Not always indicative of actual assembly failure
C. SCADA	<ul style="list-style-type: none"> • Turbine unavailability • Downtime 	<ul style="list-style-type: none"> • No maintenance activity description • Lots of data
C. Operations Planned Movements	<ul style="list-style-type: none"> • Manual acknowledgement/card swipe times for technician transfer of control on/off turbine • Estimated times for technician transfers from work plans • Repair times 	<ul style="list-style-type: none"> • Incomplete: some technicians have logged a "pick-up" from a turbine have not logged a "drop-off" from a vessel and vice-versa. • Some transfers are 'planned', but not acknowledged.

These three points are the basis upon which we define a failure going forward. However, exactly how *downtime event*, *intervention* and *corrective maintenance action* are interpreted presents an uncertainty in the failure definition. Fundamentally, the sensitivity of this interpretation is what we are investigating in sections 4 and 5.1.

3.3. Identification of Assembly-Level Failures

Assembly-level failures are identified by the following steps:

- 3.3 (i) *Downtime Catalogue* is subject to the various data selection criteria outlined in section 4.
- 3.3 (ii) If the assembly is immediately evident from the text description in *Work Procedures*, that task is labelled accordingly. Example: "Main bearing removal and replacement"
- 3.3 (iii) If the assembly is not evident from *Work Procedures*, we use the text description in *Tasks*. Tasks have two types of descriptor:
- (a) Tasks labelled with *service department* are closer to work procedure descriptors. We look for these descriptors first.
 - (b) The remaining task descriptors contain a single alarm code, corresponding to what the turbine manufacturer thinks is wrong with the turbine. These descriptors have a higher level uncertainty because they depend on accuracy of diagnosis at the beginning of the data mining process and accuracy of interpretation at the end.
- 3.3 (iv) Once tasks are labelled with an assembly, they are mapped to the enhanced taxonomy defined by Reder et al. [6]. Note that the selection of taxonomy is also a sort of "metaparameter" which has a significant impact on results - this is discussed more in section 6.1.

3.4. Key Performance Indicator Definitions

3.4.1. *Technical Availability* We define *technical availability* by the following formula [1]:

$$A_{tech} = \frac{t_{available}}{t_{available} + t_{unavailable}}, \quad (1)$$

where:

- $t_{available}$ represents the time of full and partial performance, technical standby, requested shutdown, downtime due to environment and grid.
- $t_{unavailable}$ represents the time of corrective actions, as subject to the data selection criteria of section 4.

3.4.2. Repair Time/Active Repair Time The *repair time* is similar to the mean time to repair defined by Gonzalez et al. [20]. It refers to the total duration taken to complete the entire repair process, including both active repair activities and any associated idle time or waiting time.

In contrast, *active repair time* specifically refers to the period when the actual repair work is being performed on a turbine assembly. We define active repair time as the sum of differences in pick-up/drop-off times for technicians to/from a given turbine during a given downtime event:

$$t_{activerepair} = \sum_i^I (t_{pick-up}^i - t_{drop-off}^i) \quad (2)$$

where I is the total number of work shifts throughout the period of downtime in question, $t_{drop-off}^i$ is the time a technician team is dropped-off to the turbine in the given work shift and $t_{pick-up}^i$ is the time a technician team is picked-up from the turbine in the given work shift.

4. Failure Definitions

This study defines a failure as *a turbine downtime event accompanied by an unscheduled visit to that turbine*. Often reliability analyses of wind turbines will use some variation of this definition. Carroll et al. [3], who provide the more extensive reliability analysis of offshore wind turbines in the literature, use the definition of ‘*a visit to a turbine, outside of scheduled operation, in which material is consumed*’. In the absence of any material usage data, we use turbine downtime as a qualifier for failure.

The variation in turbine-level failures initially focuses on the three sources of uncertainty highlighted in section 3.2. To elaborate:

- 4 (i) What is meant by a *downtime event*? This is a general issue addressed in one way or another by the chosen failure definition of a given wind turbine reliability analysis. It has two aspects to it.
 - (a) Some studies (e.g. Wilkinson et al. [12]) impose a lower limit on downtime for that event to be considered a failure. Failures with short down times may not be considered “real failures” as they imply no serious repair work to remedy them. In this case researchers might assume the turbine has tripped (e.g. due to grid conditions) and only requires a manual restart to bring it back online, rather than an assembly repair/replacement. We will refer to this as a *downtime limit* henceforth.
 - (b) Other studies argue that several sequential downtime events can be brought on by one failure, and that there should be some limit on the time elapsed between those events for them to be grouped under one failure. We will refer to this as a *grouping limit* henceforth.

- 4 (ii) What is meant by an *intervention*? Note that this analysis depends on *Operations Planned Movements* to scrutinise *intervention*. The applicability of the following points to the general reliability analysis depends on whether their analysis depends on a similar dataset. There are two aspects.
- For some maintenance tasks, there is a *drop-off* of a technician and no corresponding *pick-up*, and vice-versa. The question is whether to only include tasks which have both or to include tasks which have one or the other. This effectively imposes a lower limit on active repair time (see equation 2).
 - As an extension of the above point, we might impose a lower limit on active repair time. Analogously to a *downtime limit*, the *active repair time limit* might be imposed to exclude downtime events not considered "real failures" which require no serious repair work to remedy them.
- 4 (iii) What is meant by a *corrective maintenance action*? This is obvious at first glance: include all maintenance actions labelled as corrective in the database. However, it is conceivable that different studies will expand or contract the range of their included tasks subject to a number of considerations. The issues faced specifically by this dataset are listed below. Note that these are not highlighted as issues that will be faced by all reliability analyses. Since there are no set of standards around reliability data collection, each of them may or may not be applicable to the general reliability analysis.
- Opportunistic tasks.** Instead of one downtime event corresponding to one failure, one downtime event might correspond to several failures, depending on whether multiple assemblies were repaired in that time. Downtimes attributed to non-corrective works, such as annual services, sometimes contain corrective works.
 - Retrofitting.** Some retrofitting works are labelled as such in the dataset, but others are labelled as corrective.
 - Annual Services.** As above, some tasks are labelled as corrective but appear to be part of the annual service.
 - Balance of Plant (BoP) tasks.** BoP tasks cover activities associated with the infrastructure and support systems of the wind farm, excluding the assemblies that make up the wind turbine themselves. BoP jobs are either labelled as "bop - defects" or as "corrective - bop". Similar to the above, a certain amount of BoP work may be covered by contractual arrangements, and may or may not be included in failure estimates.
 - Unlabelled tasks.** Some tasks are labelled as corrective but contain no work-procedure descriptor. Others might have a vague work procedure associated with them and not a task descriptor, so similarly cannot be classified by assembly.
 - Fault Finding missions.** Some corrective tasks have the work procedure *Fault Finding* - it is unclear whether a repair was conducted directly from the fault finding mission. However, some fault finding tasks have an alarm code attached to them, and might be further categorised.

Here we define a baseline definition by the conditions:

- No downtime limit;
- No grouping limit;
- No opportunistic maintenance included (each downtime event is one failure);
- All tasks that are recorded as corrective are included. This includes some tasks which might otherwise be recorded as retrofitting and annual service.

5. Results

5.1. Turbine-Level Failures

The baseline estimate comes in at 9.06 failures per turbine per year. Figure 1 explores the sensitivity of that figure to the various points outlined in section 4. Figure 1(a) shows the failure rate estimate falling sharply from the baseline with increasing downtime limit up until around 10 hours, after which the decline starts to slow. The limit of the x-axis corresponds to the most extreme downtime limit in the literature of 72 hours [1]. At the more reasonable limit of 1 hour, the failure rate estimate reaches 8.5 failures per turbine per year, a 6% reduction on the baseline. However, it is evident from figure 1 that dropping these failures does not incur a significant increase in availability. Figure 1 (c) shows the dropping failure rate estimate and rising technical availability up to a downtime limit of 5 hours. Only slight variations in technical availability are evident in this range.

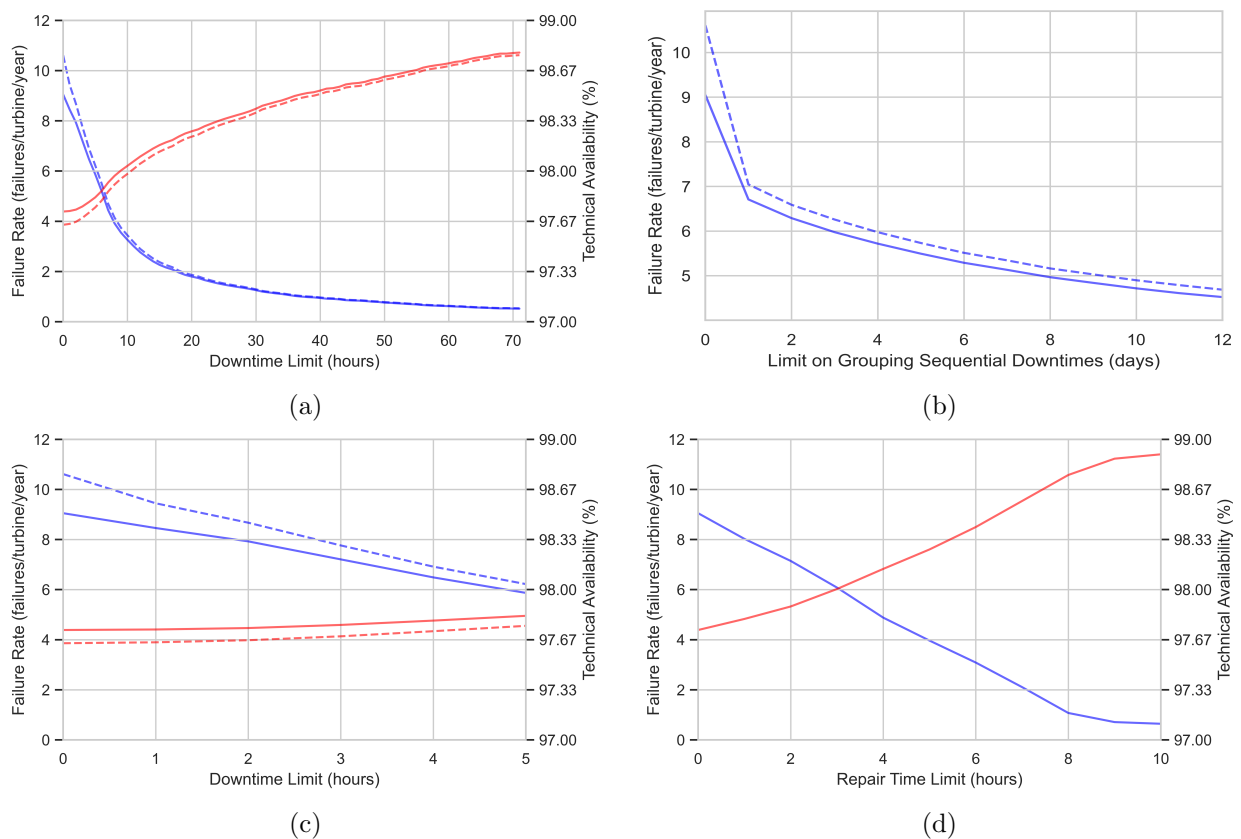


Figure 1: Sensitivity of baseline failure rate (blue lines) to (a) downtime limit, (b) grouping limit and (d) repair limit. (c) Also shows the downtime limit from 0 to 5 hours on the x-axis. Red lines show the corresponding technical availability. Solid lines include only downtime events with *both* manually acknowledged drop-off and pick-up; dashed lines include downtime events with *either* a manually acknowledged drop-off or pick-up.

Figure 1 (b) shows a similarly sharp fall-off from the baseline case with increasing grouping limit up until the 24 hour mark, after which the curve flattens. At 24 hours, the failure estimate reaches 6.8 failures per turbine per year. Note that the availability estimate remains constant independent of the grouping limit.

Repair time limit is explored in figure 1 (d). This shows an approximately linear relationship between failure rate estimates and repair time limits. At 1 hour, the baseline drops to 8.03; at

two hours to 7.15. The effect of increasing repair time limit on technical availability is more significant, implying more care should be taken in imposing even small repair time limits.

Figures 1 (a), (b), and (c) address the question: *what is meant by an intervention?*. Including zero repair time jobs (dashed lines) increases the baseline failure rate estimates to 10.77. This disparity decreases with increasing downtime limit.

Figure 2 explores the question: *what is meant by 'corrective maintenance'*. Essentially the baseline failure definition can be expanded or contracted based on additional data selection criteria. Of the data selection criteria explored, the failure rate ranges from 7.38 to 12.12. Inclusion of opportunistic jobs increases the failure rate to 10.77 failures per turbine per year. This has a particularly significant effect on technical availability, reducing the baseline estimate by 1.38%. Note that this availability estimate contains all downtimes where a corrective action was carried out, even if the majority of the downtime was due to (e.g.) an annual service. BoP jobs increase the baseline to 9.38 failures. However, the baseline estimate can also be decreased via plausible data selection measures. Filtering 'no-assembly' and 'fault finding' missions have a similarly significant effect to the inclusion of opportunistic jobs, reducing the baseline to 7.38 and 7.45 respectively. Filtering out jobs which could be alternatively labelled 'retrofit' or 'annual service' reduces the baseline to 7.79. The technical availability reduction in all of these scenarios is similar, at around 2%.

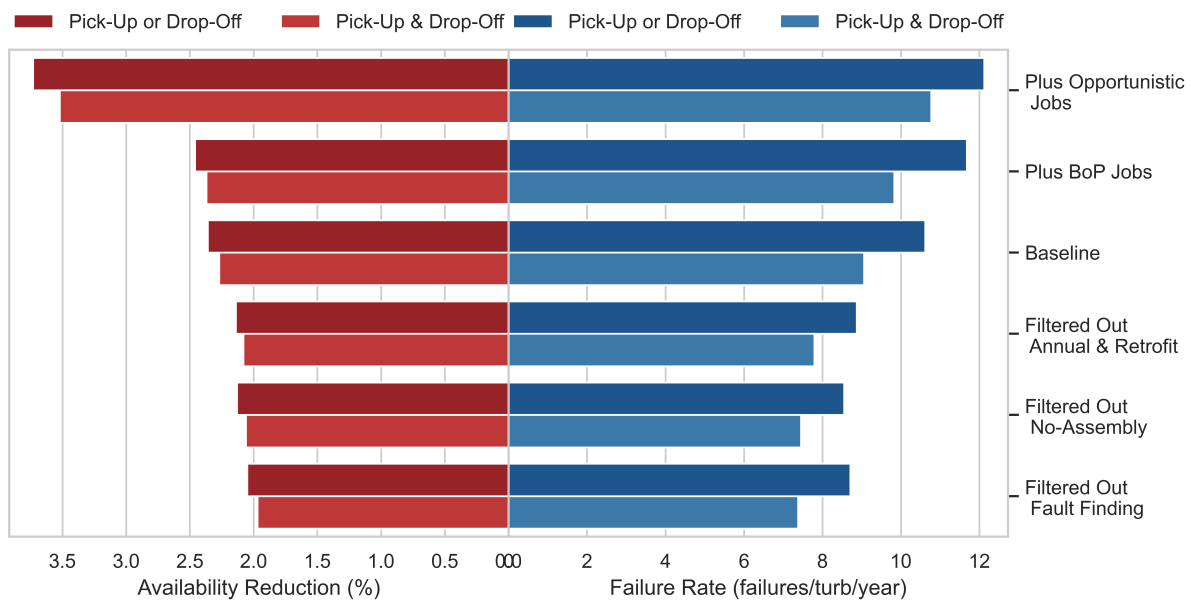


Figure 2: Bar chart exploring the sensitivity of the failure rate estimate (blue) and technical availability drop (red) to various data selection criteria.

5.2. Assembly-Level Failures

The baseline for assembly level failures is altered to reflect the results of section 5.1. Namely, the following is added to the baseline definition:

- 5.2 (i) A downtime limit of 1 hour.
- 5.2 (ii) Jobs which contain task descriptions in line with retrofits and annual services are labelled as such.
- 5.2 (iii) 'No assembly' jobs are filtered.

Note, the baseline is re-defined by the above parameters so that the additional failures from different scenarios can be made obvious. It is not redefined to capture the 'best' failure definition.

Figure 3 shows the baseline estimates of assembly-level failures via the black bars. Most failures arise from the converter and the wind turbine cooling system, followed by the gearbox, yaw system, blades and hydraulic group. Note that, assemblies are defined by function. This means that assemblies like (e.g.) the cooling system, hydraulic group and drive train bearings might otherwise be categorised under different assemblies. Opportunistic jobs are shown by the blue bars. Additional BoP jobs are shown by red bars - they are restricted to the service crane and lift. Retrofit are shown by green bars - they have the most evident effect on the converters and cooling system, assemblies which are retrofitted regularly.

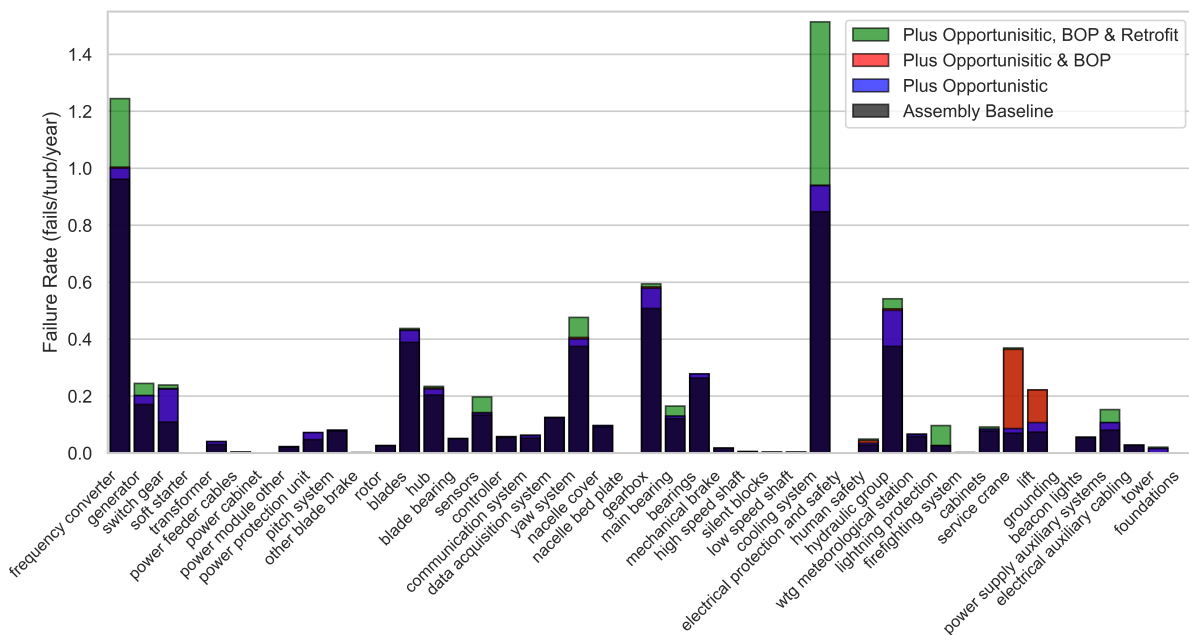


Figure 3: Assembly-level failure rates under different data selection criteria, shown in units of failures per turbine per year.

6. Discussion

6.1. Reflections on Methodology for Wind Turbine Reliability Analysis

This paper is unusual as it effectively presents the messy data pre-processing stage of a wind turbine reliability analysis. A publication of this sort is valuable to the research community as it exposes the uncertainty surrounding reliability analysis of wind turbines. From this one dataset, results in section 5 show derived failure rates ranging from below 1 failure per turbine per year to over 10 failures per turbine per year using failure definitions which have previously been used in the literature. Main sources of uncertainty for the failure rate estimate come (1) data-preprocessing and (2) any limits placed on potential failures. Figure 2 explores the range of uncertainty in point (1); failure rate estimates range from 7.38 to 12.12 depending on which tasks are included/excluded. Figure 1 explores the range of uncertainty in point (2); the *downtime limit* reduces the baseline failure rate estimate by approximately 0.5 failures per turbine per year with every hour added to the limit, up until around 10 hours. The *active repair time limit* reduces the baseline failure rate approximately 1 failures per turbine per year with every hour added to the limit, up until around 8 hours. The most significant reduction in the *grouping limit*

is in the first 24 hours, where the baseline estimate is reduced to 6.8 failures per turbine per year.

A lack of any standard failure definition therefore presents a significant epistemic uncertainty to any wind turbine reliability analysis. On top of this, a lack of any standards on data collection and processing (indeed a lack even of discussion on data processing) introduces further uncertainty. In fact, the uncertainty revealed via the above results indicate a standard failure definition would be insufficient, and that there is a need for a recommended practice on data collection.

As reliability analyses are rare and valuable in the wind industry, one has been undertaken on the available dataset. There are additional uncertainties which affect the results and have not been addressed here and are worth addressing. Firstly, there is uncertainty around data interpretation. Another group of researchers may have produced different results for assembly-level failures, as many of them depend on alarm codes to ascribe them to a specific assembly. Given the dataset does not record the complete history of alarm codes, but simply the most likely culprit for the fault, this is something of a risky tactic. Secondly, there is no standard taxonomy in the wind industry. The taxonomy therefore becomes a sort of hyper-parameter which effects how the data is interpreted. A similar study could be undertaken to map the work procedures to other taxonomies used in the literature, most notably the RDS-PP taxonomy [2].

Secondly, the dataset upon which this study is based is from a single offshore wind farm. As a consequence: (a) failure rates are from one manufacturer and turbine model, and therefore represent a small subsection of turbines available on the market; (b) the dataset only covers a small portion of the farm's lifetime, and may mis-represent the average failure rates of assemblies over the turbine's lifetime.

Thirdly, the scenario 'Filtered out annual service and retrofit' is largely a data quality issue. According to the operator, most tasks which are recorded as corrective, but whose work procedure imply a retrofit or annual service, are corrective in nature. This is why they were included in the 'baseline' calculation. Often when a turbine fails, and the solution is to retrofit it. With increasing take-up of better data standards on the industry, these kind of interpretation issues will be obsolete. As it stands, at least for this data set, this is another instance which requires manual interpretation, and could be taken either way depending on the mindset of the analyst.

6.2. Future Analysis based on this Dataset

There are several future works which could provide value to the research community. Most immediately, value could also be extracted by replicating Carrol et al.'s reliability study for offshore turbines [3]. Failure rates are included here, but the study could be extended by categorising failures by severity, calculating repair times and number of technicians required to repair. Since SCADA data is available, downtime and lost production could also be included. These metrics could also be calculated for non-corrective works, records of which are also sparse in the literature.

Second, we analysed the effects of a *downtime limit*, *grouping limit*, *active repair time limit* and inclusion of different maintenance works separately. It is foreseeable, however, that a reasonable failure definition (i.e. reasonable for a particular use-case) would incorporate multiple of these elements. It would therefore be interesting to see the joint effect of these factors.

Third, a methodology could be employed to assess uncertainty both inherent in the data itself and in the calculated metrics. The current authors have employed Bayesian techniques to do so in the field of wind industry in previous studies [19, 21]. Both of these studies could be improved by retroactively including the work procedure data upon which this study relies so heavily. On the other hand, novel Bayesian techniques could be used to address uncertainty in the data collection process itself.

7. Conclusions

This study presents a reliability analysis of wind turbines at the turbine level and the assembly level. Its aim is to present failure data which is rarely presented and therefore valuable within the research community. However, it also aims to lay bare the considerable uncertainty associated with data-preprocessing of reliability data. We do so by exploring the sensitivity of failure rate estimates to different failure definitions and data selection criteria. The baseline failure definition used is 'A turbine downtime event accompanied by an unscheduled visit to that turbine'. Different interpretations of *downtime event* were explored by imposing a lower limit on the downtime of an event for it to be considered a failure; a similar lower limit on repair time; and a limit on the amount of time allowed to elapse between sequential downtime events at the same turbine to be grouped into one failure. The baseline failure rate estimate of 9.06 failures per turbine per year showed a considerable sensitivity to all of these factors. Including opportunistic jobs in the estimate increases the baseline by 19%, including corrective balance of plant jobs by 8%. Filtering the corrective maintenance actions to exclude those which could also be interpreted as retrofitting or annual service activities reduces the baseline by 16%. Filtering out jobs which couldn't be fit into the assembly taxonomy reduced the baseline by 18%, and filtering 'fault finding' missions reduced the baseline by 19%. Assembly failure rates showed high values for the frequency converter and cooling system, which were radically increased by including retrofit jobs.

References

- [1] Pfaffel S, Faulstich S and Rohrig K 2017 Performance and reliability of wind turbines: A review *Energies* **10** 1904
- [2] Cevasco D, Kolios A and Koukoura S 2021 Reliability, availability, maintainability data review for the identification of trends in offshore wind energy application *Renewable and Sustainable Energy Reviews* **136** 110414
- [3] Carroll J, McDonald A and McMillan D 2019 Failure rate, repair time and unscheduled O&M cost analysis of offshore wind turbines *Wind Energy* **19** 1107–1119
- [4] 2017 Sparta Portfolio Review 2016/17
- [5] Leahy K, Gallagher C, O'Donovan P and O'Sullivan D T J 2020 Issues with data quality for wind turbine condition monitoring and reliability analyses *Ocean Engineering* **209** 107381
- [6] Reder M D, Gonzalez E and Melero J J 2016 Wind Turbine Failures - Tackling current Problems in Failure Data Analysis *Journal of Physics: Conference Series* **753**
- [7] Hahn B, Welte T, Faulstich S, Bangalore P, Bousson C, Harrison K, Miguelanez-Martin E, O'Connor F, Pettersson L, Soraghan C *et al.* 2017 Recommended practices for wind farm data collection and reliability assessment for o&m optimization *Energy Procedia* **137** 358–365
- [8] Seyr H and Muskulus M 2017 Interaction of repair time distributions with a weather model *29th International Congress on Condition Monitoring and Diagnostics Engineering Management (COMADEM 2016)*
- [9] Scheu M N, Kolios A, Fischer T and Brennan F 2017 Influence of statistical uncertainty of component reliability estimations on offshore wind farm availability *Reliability Engineering and System Safety* **168**
- [10] Dao C D, Kazemtabrizi B and Crabtree C J 2020 Offshore wind turbine reliability and operational simulation under uncertainties *Wind Energy* **23**
- [11] Martin R, Lazakis I, Barbouchi S and Johanning L 2016 Sensitivity analysis of offshore wind farm operation and maintenance cost and availability *Renewable Energy* **85** 1226–1236
- [12] Wilkinson M, Hendriks B, Spinato F, Gomez E, Bulacio H and Roca J e a Measuring wind turbine reliability-results of the reliawind project
- [13] Danish Standards Foundation 2020 *A Guide to RDS - Reference Designation Systems. TAG Numbers for Systems in Accordance with the ISO/IEC 81346 Standard Series*
- [14] NERC 2018 *Wind Turbine Generation: Data Reporting Instructions*
- [15] ISO 2016 *ISO 14224: 2016 Petroleum, Petrochemical and Natural Gas Industries—Collection and Exchange of Reliability and Maintenance Data for Equipment*
- [16] FGW ZEUS 2013 *Technical Guidelines for Power generating Units, Part 7: Maintenance of power plants for renewable energy, Category D2: State EventCause Code for power generating units*
- [17] International Electrotechnical Commission (IEC) 2010 *IEC 61400-26-1: Time Based Availability for Wind Turbines*

- [18] Anderson F, Dawid R, García Cava D and McMillan D 2021 Operational metrics for an offshore wind farm their relation to turbine access restrictions and position in the array *Journal of Physics: Conference Series* **012002**
- [19] Anderson F, McMillan D, Dawid R and Garcia Cava D 2023 A bayesian hierarchical assessment of night shift working for offshore wind farms *Wind Energy* **26** 402–421
- [20] Gonzalez E, Nanos E, Seyr H, Valdecabres L, Yürüşen N Y, Smolka U, Muskulus M and Melero J J 2017 Key performance indicators for wind farm operation and maintenance *Energy Procedia* **137** 559–570
- [21] Anderson F, Dawid R, McMillan D and García Cava D 2023 A bayesian reliability analysis exploring the effect of scheduled maintenance on wind turbine time-to-failure *Manuscript under review*