A Quantitative Framework for Diagnosing Energy Losses in Wind Turbine Operational Logs: A Single-Turbine Case Study

Aditya Rajiv Ratnam

National Academy for Learning, Bengaluru

Abstract

Wind turbine operators routinely record daily generation alongside categorized downtime, yet this operational log data is typically reviewed only to confirm that a turbine generated electricity on a given day, without a systematic accounting of why its output fell short of potential. This study presents a reproducible framework for decomposing a wind turbine's total available time into five categories (generating, grid-unavailable, mechanically broken down, under maintenance, and idle for insufficient wind), classifying the documented cause of lost time from free-text operational remarks, and statistically testing whether the resulting patterns in downtime timing, cause, and season are distinguishable from random variation. The framework is demonstrated on 1,614 consecutive days (1 April 2022 to 31 August 2026) of daily operational logs from a single 250 kW wind turbine generator in Tamil Nadu, India. Both mechanical breakdown and grid-related outages showed inter-event timing inconsistent with a homogeneous Poisson process (Monte Carlo-corrected $p = 0.0005$), suggesting temporal clustering, though the evidence is confounded by seasonality and does not, by itself, establish that individual events are causally dependent. The near-total drop in generation observed January through March each year was statistically large (Kruskal–Wallis $\varepsilon^2 = 0.77$) and coincided with the turbine being overwhelmingly classified as Lull rather than Grid-Down or Breakdown. Using month-specific generation rates rather than a single site-wide average reversed the ranking of the two largest documented causes by hours lost: Weather-related hours exceed Grid-infrastructure hours, but Grid-infrastructure losses are worth more in estimated annual revenue, because Weather losses concentrate in a lower-generation season. However, the single largest category of estimated generation opportunity loss overall was downtime with no recorded cause, exceeding any specific documented cause. The analysis pipeline, released as open-source, is potentially applicable to other turbines recording comparable fields, subject to validation of category definitions and data quality.

Keywords: wind turbine, reliability, grid curtailment, renewable energy, downtime analysis, operational log analysis, Weibull distribution, nonparametric statistics, Kruskal-Wallis, Kolmogorov–Smirnov

1. Introduction

Global installed wind capacity has grown by more than a factor of one hundred since the late 1990s, and wind energy is expected to supply a large share of future decarbonized electricity systems (International Renewable Energy Agency [IRENA], n.d.). India's wind sector is concentrated disproportionately in a small number of states, of which Tamil Nadu holds the largest installed wind capacity in the country, driven by strong and highly seasonal winds associated with the southwest monsoon. Because Tamil Nadu's windy sites are so strongly tied to the monsoon, wind farm output in the state can swing from near its rated capacity to

almost zero within the same year, a pattern that complicates both grid planning and the interpretation of any individual turbine's operating record (Bastin et al., 2023).

Wind turbine operators typically maintain a daily or hourly operating log recording generation, run hours, and time lost to various causes. Logs of this kind are routinely maintained for commercial turbines and underpin formal availability standards such as IEC 61400-26-1 (International Electrotechnical Commission [IEC], 2019). In practice, however, these logs are usually reviewed only to confirm that a turbine generated electricity on a given day, with comparatively little systematic analysis of how time is actually being lost, whether that lost time follows a discernible statistical pattern, or how its documented cause compares across categories. Although low capacity factors and unreliable grid delivery are well-documented barriers to renewable energy utilization in general (George & Banerjee, 2009), such logs are less commonly analyzed as standalone datasets unless a utility or manufacturer is formally required to report on them.

This study's main contribution is a reproducible, general-purpose analytical framework that decomposes a wind turbine's operating record into distinct time categories, statistically tests the pattern of losses across those categories, and estimates the generation opportunity loss associated with each documented cause of loss, using only the daily reading log most operators already keep. Unlike a conventional wind turbine reliability study, which typically models the time to failure of individual mechanical or electrical subsystems from SCADA data (Dao et al., 2025), this framework works directly from the site-level daily reading log already produced by the turbine's monitoring system, without requiring subsystem-level sensor data. The single-turbine case study demonstrates the framework rather than establishing a general finding about Tamil Nadu wind turbines.

This study has three objectives: first, to classify a turbine's total available time into five categories reflecting generation, grid unavailability, mechanical breakdown, maintenance, and insufficient wind, and to further attribute Grid-Down and Breakdown time to a documented cause using the turbine's own free-text operational remarks; second, to test whether the resulting patterns in downtime timing, seasonality, and cause are statistically distinguishable from random variation; and third, to demonstrate this framework on a single turbine's complete four-and-a-half-year operating history as a worked example.

2. Site and Data Description

The dataset comprises the daily operational reading log of a single 250 kW wind turbine generator in Tamil Nadu, India, provided by the turbine's owning company from the site's routine daily reading records. Site- and equipment-identifying details beyond those in Table 1 have been withheld from this study. The log spans 1,614 consecutive calendar days, from 1 April 2022 through 31 August 2026, and every calendar month in that range is represented.

Table 1 summarizes the site and the fields recorded in the log. Each daily row records the turbine's generation in kWh (from two generator readings, G1 and G2) along with five time categories that are together intended to account for the full 24 hours of that day: Run Hours, during which the turbine was generating; Grid-Down Hours, during which the turbine was available but the grid connection was not; Breakdown Hours, during which the turbine itself was non-operational; Maintenance Hours, during which the turbine was deliberately taken offline for servicing; and Lull Hours, during which the turbine was available and

grid-connected but wind speed was insufficient to generate. A free-text remarks field, populated on 361 of the 1,614 days (22.4%), records the operator's account of unusual events, most commonly grid-side failures such as line faults, feeder trips, and storm damage.

Field	Value
Turbine type	Single wind turbine generator, manufactured by SEPC
Rated capacity	250 kW
Location	Tamil Nadu, India
Data source	Daily reading log from site owner records
Dataset span	1 April 2022 to 31 August 2026 (1,614 consecutive days, no missing months)
Recorded fields	G1/G2 generation (kWh); Run, Grid-Down, Breakdown, Maintenance, and Lull hours; EB Export/Import/Net Export; RKVAh; RKVAR %; M/A % and G/A %; free-text remarks

Table 1*. Site and dataset summary.*

3. Methodology

3.1 Data Processing Pipeline

The site log is maintained as a set of monthly worksheets, one per calendar month, with inconsistent sheet-naming conventions accumulated over more than four years of manual record-keeping (for example, "April2022", "April 2026", and "JULY2025" all appear as sheet names for different months). A parsing module reads every worksheet, normalizes the sheet name to a calendar year and month regardless of its exact formatting, and extracts each day's twenty recorded fields into a single table indexed by date. The resulting daily table, and every result in this study, was generated by an open-source Python pipeline built on pandas, NumPy, SciPy, and Matplotlib (Ratnam, 2026).

Two non-numeric data-entry artifacts were found during parsing: a literal non-numeric character in place of the G1 generation reading, on two consecutive days in March 2024. A full scan of every numeric field across all 53 monthly worksheets confirmed these two cells are the only non-numeric values anywhere in the dataset's more than 32,000 extracted data points; the dataset is otherwise clean. Rather than coercing the two affected cells to zero without justification, the surrounding fields on both days were checked: Run Hours was recorded as 0 and Grid-Down Hours as the full 24 hours on both days, meaning the turbine could not have generated regardless of the specific value the corrupted field held. Therefore, setting G1 to zero on these two days is justified.

A separate check cross-referenced the number of parsed daily rows against the number of calendar days spanned by the stated date range and found a one-day shortfall. The row for 26 June 2026 had its day-of-month field entered as 0 rather than 26 in the source spreadsheet. This was corrected at the source by setting the field to 26. All figures, tables, and statistics reported in this study reflect the corrected 1,614-day dataset.

Summing the five time categories against the 38,736 possible hours in the dataset (Table 2, Section 3.2) leaves a small residual of about 221 hours (0.57%). Rather than attributing this to rounding without evidence, the residual was investigated directly. It decomposes into three distinct, identifiable sources. First, eight days have all five categories recorded as exactly zero, consistent with a missing daily record rather than a genuine zero-hour condition (192

hours, 0.50% of the dataset). Second, two days (29 and 30 April 2025) have both Run Hours and Lull Hours recorded as the full 24 hours on the same day. This double-counting indicates a duplicate or overlapping entry rather than two genuinely co-occurring categories. Third, one day (30 May 2026) has only 4 of 24 hours categorized at all, an apparently truncated entry. Once these eleven days are set aside, the remaining 141 days with a nonzero residual show a median gap of only 0.4 hours, consistent with ordinary rounding in partial-hour recording. These residual days sum to just 57.1 hours in total, only 0.15% of the dataset. Therefore, the majority of the original 221-hour residual is attributable to a small number of specific, identifiable data issues, not to rounding spread evenly across the dataset.

3.2 Loss Categories and Cause Classification

The five time categories recorded in the log (Section 2) are the site's own operational categories. They map loosely onto the generic categories defined by IEC 61400-26-1, the formal standard for time-based wind turbine availability reporting (International Electrotechnical Commission [IEC], 2019): Run Hours corresponds approximately to that standard's Full Performance category, Grid-Down to its External Environment or Grid Unavailable categories, Breakdown to Unscheduled Maintenance, and Maintenance to Scheduled Maintenance, while Lull Hours has no direct IEC equivalent, since the standard's categories describe turbine and grid status rather than wind resource availability. This study uses the site's own category definitions throughout rather than remapping the data into the standard's scheme, and this correspondence is approximate rather than exact.

Grid-Down and Breakdown hours represent potential generation opportunity loss, unlike Lull hours, which reflect an absent wind resource rather than a turbine fault. These two categories were further attributed to a cause wherever the day's remark field was populated. A rule-based keyword classifier sorts each remark into one of six categories: weather (rain, lightning, storm); grid infrastructure (EB line, feeder, transformer, pole, substation); vegetation (trees fouling lines); maintenance or rework; mechanical or electrical fault; and low wind. Any remaining populated remark falls into a seventh, catch-all "other/unclassified" category. Any day with no remark at all is left explicitly unclassified (labeled "unattributed" in Section 4 to distinguish a documentation gap from a cause). No remark in this dataset matched the vegetation rule, so it does not appear as a category anywhere in Section 4. This classifier is keyword-based rather than a trained model since no labeled dataset of cause categories exists against which to train or validate one. The six categories and their keyword rules were defined by inspecting the vocabulary actually used in the remarks. Only 22.4% of days in the dataset have any remark at all, and those days are not representative of total lost time. Days with remarks account for only 46.5% of total Grid-Down plus Breakdown hours, which means the cause classification characterizes a minority of lost hours.

3.3 Statistical Methods

3.3.1 Event extraction, MTBF, and MTTR

Downtime in the Breakdown and Grid-Down categories was converted from daily hour totals into discrete events. A run of one or more consecutive calendar days on which the category's hours were greater than zero was treated as a single continuous event, rather than as several independent one-day events. Breakdown and Grid-Down events were extracted as two separate streams; three days in the dataset (0.19%) have a nonzero value in both categories

simultaneously, each for a small fraction of the day (at most a few hours), and each such day contributes to one event in each stream rather than to a combined category. This daily-resolution definition does not capture within-day start and stop times, so an event beginning late on one day and continuing into the next is treated identically to one that spans the entire first day. Section 7 (Limitations) discusses the resulting sensitivity of the reliability estimates to this choice. Mean Time To Repair (MTTR) was computed as the mean duration of these events; Mean Time Between Failures (MTBF) was computed as the mean number of days between the end of one event and the start of the next.

3.3.2 Weibull fit and bootstrap confidence intervals

A two-parameter Weibull distribution was fitted, by maximum likelihood, to the gaps between successive downtime events, following the approach used in wind turbine reliability studies that model time-between-failures directly from operational data (Dao et al., 2025). Weibull fitting only describes the shape of the gaps between events; it does not model the underlying failure mechanism. This is a different use of Weibull than either of its two familiar roles in wind energy: modeling the distribution of wind speeds at a site, or modeling how long a single aging component lasts before failing (the bathtub-curve approach). Here, the system resets after every event, so the shape parameter k describes the pattern of gaps between repeated events, not a wind-speed profile or one component's lifespan. A shape parameter below one indicates unevenly distributed gaps, with many short gaps punctuated by occasional long quiet stretches, consistent with failures clustering in bursts. A shape parameter near one is consistent with a memoryless, randomly timed process. A shape parameter above one indicates more evenly spaced, regular recurrence. The maximum likelihood equations for the Weibull shape and scale parameters have no closed-form solution and their estimators are known to carry first-order bias, with sampling distributions that are not always well approximated by a normal distribution (Truong et al., 2025). The mean time between failures is estimated directly from a small, skewed sample of gaps, for which a normal approximation is similarly unreliable. Therefore, a 95% confidence interval was placed on both the Weibull shape parameter and the mean time between failures using a percentile bootstrap (2,000 resamples of the observed inter-event gaps, with replacement), rather than relying on an asymptotic normal approximation.

3.3.3 Kolmogorov–Smirnov test with a Monte Carlo-corrected p-value

As a complementary check on the Weibull-based clustering finding, a Kolmogorov-Smirnov goodness-of-fit test compared the observed inter-failure gaps against an exponential distribution. This is the pattern the gaps would follow if downtime occurred as a homogeneous Poisson process, meaning independently and at a constant average rate. The standard way of calculating a p-value for this test assumes the exponential distribution's rate was chosen in advance, not estimated from the very data being tested. Since the rate here was estimated from the same data, that standard p-value cannot be trusted. This study therefore also reports a more reliable Monte Carlo-corrected p-value, computed as follows. A synthetic sample of the same size is repeatedly drawn from an exponential distribution with the fitted rate, the rate is then refit to each synthetic sample in turn, and its own Kolmogorov–Smirnov statistic is computed. This procedure is repeated over 2,000 such simulations. The corrected p-value is calculated as $(b + 1) / (n + 1)$, where b is the number of simulations that produced a statistic at least as extreme as the one actually observed and n is the number of simulations. This is the standard Monte Carlo estimator for this situation, used instead of the naive b / n,

because a count of zero extreme simulations would not prove that the true p-value is exactly zero, only that it is smaller than this simulation can measure. The coefficient of variation (CV) of the gaps was also calculated alongside both p-values, as a second way of measuring the same effect. CV equals 1 for a purely random process, and rises above 1 when events are more clustered ("over-dispersed") than random chance would produce. Rejecting the random, constant-rate hypothesis is evidence against purely random timing, but a failure rate that simply varies smoothly with the seasons, with no real event-to-event clustering at all, could produce this same result. So this finding is read as evidence of clustering in a descriptive sense, not as proof that individual failures are actually linked to each other. Telling these two explanations apart would require a more advanced model that accounts for seasonality directly, which is left for future work (Section 8).

3.3.4 Chi-square test of cause versus season, with robustness steps

Each remarked day's classified cause was cross-tabulated against season, counting how many days of each cause fell in each season (Winter: December to February; Summer: March to May; Monsoon: June to September; Post-Monsoon: October to November), using the Indian Meteorological Department's (IMD) standard calendar. A chi-square test of independence was then applied to this table. This test checks whether the pattern of causes across seasons looks different from what pure chance would produce, which would mean cause of loss is genuinely associated with season. One problem with this table is that several of its cells had very few days in them, and some had zero. Chi-square tests are normally considered reliable only when most cells have an expected count of at least five; with six cause categories spread across four seasons, several cells fell short of that.

Two further steps were therefore taken. First, the six cause categories were merged into a coarser five-group scheme (Wind/resource, Grid, Equipment, Weather, Other/unknown). This raised the share of cells meeting the standard "at least five" rule of thumb (McDonald, 2014), making the test more trustworthy. Second, a permutation test was run on the merged table. This method sidesteps the "at least five" rule entirely: the computer randomly reshuffles which season each of the 361 remarked days is assigned to, 2,000 times, recalculating the chi-square statistic after each shuffle. Since this shuffling breaks any real link between cause and season, comparing the real, unshuffled result against these 2,000 randomized ones shows how likely the real pattern is to have occurred by chance alone. Finally, Cramer's V was calculated for both the original and merged tables. This is a standard measure of how strong the association between cause and season is, on a scale from 0 (no association at all) to 1 (a perfect one-to-one relationship).

3.3.5 Kruskal–Wallis test for seasonality

The apparent seasonal pattern in daily generation and daily lull hours was tested using the Kruskal–Wallis test, the non-parametric equivalent of one-way ANOVA (analysis of variance), applied across the twelve calendar months. This test was preferred over ANOVA because daily generation and daily lull hours are bounded and strongly skewed rather than approximately normally distributed. Epsilon-squared, an effect-size measure for the Kruskal–Wallis test bounded between 0 (no effect) and 1 (groups fully separated), was reported alongside the p-value, since with over 1,600 daily observations, even a small real effect would produce a very small p-value on its own. This test, like the clustering tests above, assumes the daily observations being compared are independent. This assumption is

questionable because daily generation and lull hours from the same site across multiple years are not necessarily fully independent, since weather and operating conditions can be serially correlated from one day to the next. This independence assumption is not verified in this study and is noted as a limitation in Section 7.

3.3.6 Economic sensitivity of loss causes, using month-specific rates

To translate the cause-classified loss hours into a decision-relevant estimate, each cause category's estimated generation opportunity loss was approximated as follows. For every remarked day, that day's Grid-Down plus Breakdown hours were multiplied by that calendar month's own average generation rate per run-hour, rather than a single rate averaged across the whole dataset, and these day-level values were summed. This is a deliberate refinement over a blanket site-wide average: a cause concentrated in a low-generation month should not be valued at the same rate as one concentrated in a high-generation month. The estimate was then converted to an estimated annual revenue figure using the site's confirmed tariff of ₹3.39 per kWh, which applies to every day in the observation period. Annualized estimates were obtained by dividing each cause category's total estimated generation opportunity loss over the 1,614-day observation period by the observation period expressed in years (calendar days divided by 365.25, approximately 4.42 years), rather than by a fixed four-year divisor, so that Table 9 is fully reproducible from this Methods description alone. Even with month-specific rates and a confirmed tariff, this proxy generation rate remains an approximation, not an audited loss figure since actual generation during any specific downed hour would depend on the wind conditions at that exact time, which the dataset does not record. This approximation is optimistic for hours lost during severe weather, when little generation may have been possible regardless, and conservative for hours lost during otherwise-favorable conditions. The resulting figures are presented as a tool for prioritizing causes against one another, not as a guaranteed recoverable amount.

4. Results

4.1 Loss Decomposition

Table 2 summarizes how the dataset's 38,736 possible hours (1,614 days times 24 hours) are allocated across the five categories.

Category	Hours	% of possible time
Run (generating)	14,848.7	38.33%
Lull (low wind)	21,111.2	54.50%
Grid-Down	2,262.1	5.84%
Breakdown	235.95	0.61%
Maintenance	57.0	0.15%
Total accounted	**38,514.9**	**99.43%**

***Table 2**. Allocation of the dataset's 38,736 hours across the five recorded time categories.*

Figure 1 shows the monthly breakdown of the five time categories across the full dataset. Lull hours dominate the total time allocation (54.50% of all possible time; Table 2), exceeding Run hours (38.33%), while Grid-Down and Breakdown together account for only 6.45% of

possible time. This is a small fraction of total time, but a potentially consequential source of generation opportunity loss, since unlike Lull hours, which reflect an absent resource, these hours reflect operational or grid-side unavailability.

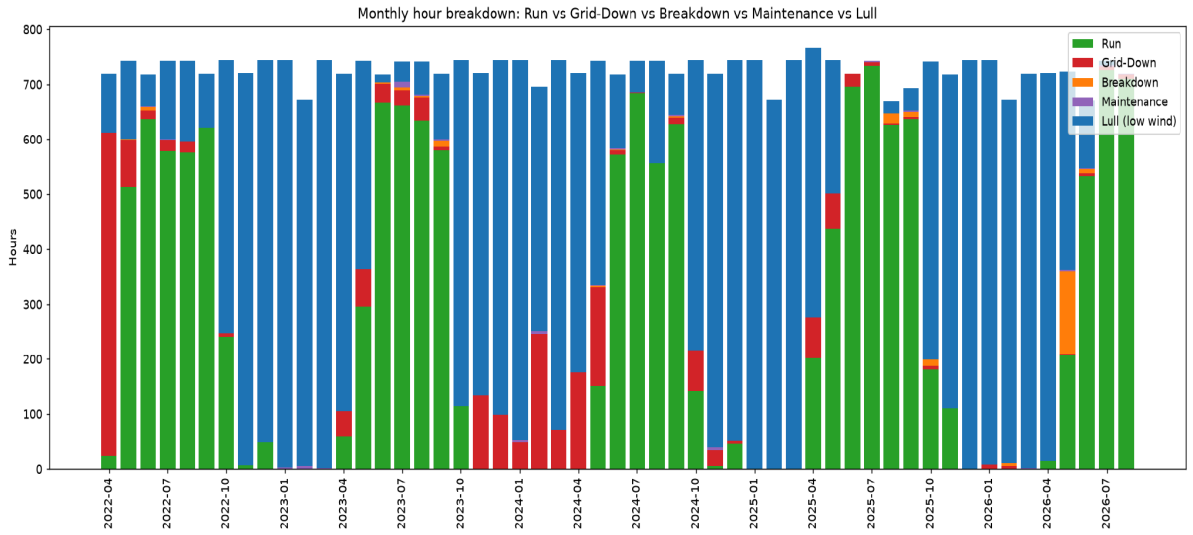

***Figure 1**. Monthly hour allocation across the five recorded time categories (Run, Grid-Down, Breakdown, Maintenance, and Lull, in hours) for the single 250 kW turbine analyzed in this study, April 2022 to August 2026 (1,614 days total).*

Figure 2 ranks these Grid-Down and Breakdown hours by their classified cause, using the keyword-based remark classification described in Section 3.2. The single largest category by a wide margin is downtime with no remark recorded at all, exceeding the combined total of every specifically documented cause (Weather, Grid infrastructure, Other/unclassified, Mechanical/electrical fault, Low wind, and Maintenance/rework). Section 4.6 reports the annualized generation-opportunity-loss and revenue estimate behind this ranking.

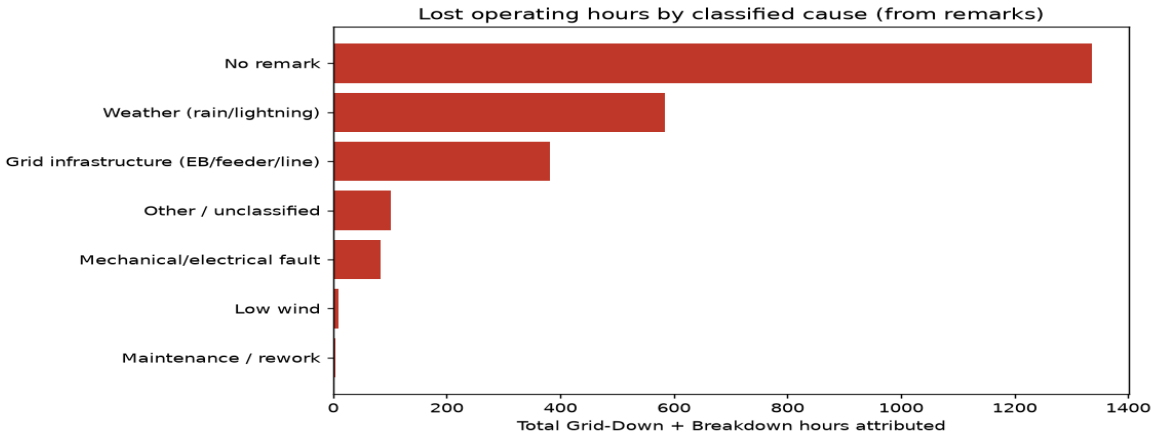

***Figure 2**. Total lost operating hours (Grid-Down plus Breakdown) attributed to each classified cause of loss, including a "No remark" category for days with no remark at all, kept separate from the "Other/unclassified" category for remarked-but-unclassified days.*

4.2 Seasonal Pattern

Table 3 and Figure 3 summarize generation and lull hours by calendar month, aggregated across all years in the dataset. January, February, and March show exactly zero recorded generation in every year covered by the dataset, with 90 to 98% of possible hours in each of

those months classified as Lull, meaning the turbine remained available and grid-connected but wind speed was insufficient. Generation resumes in April, rises sharply through the monsoon months of June through September, and declines again from October. Section 5 discusses this pattern alongside published regional wind-resource literature.

Month	Total generation (kWh)	Avg. daily generation (kWh)	Lull hours	Lull % of possible
January	0	0.0	2,912.35	97.9%
February	0	0.0	2,447.00	90.2%
March	0	0.0	2,877.05	96.7%
April	7,720	51.5	2,462.75	68.4%
May	121,392	783.2	1,532.75	41.2%
June	320,044	2,133.6	328.10	9.1%
July	370,520	2,390.5	242.20	6.5%
August	315,393	2,034.8	414.30	11.1%
September	236,374	1,969.8	334.20	11.6%
October	50,108	404.1	2,196.00	73.8%
November	5,651	47.1	2,587.05	89.8%
December	5,533	44.6	2,777.40	93.3%

Table 3. *Generation and lull hours by calendar month, aggregated across April 2022 to August 2026.*

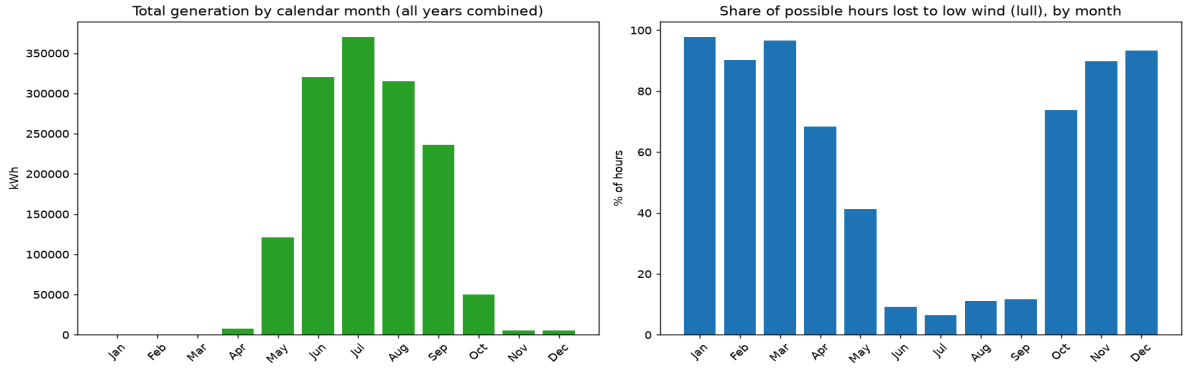

Figure 3. *Total generation (kWh, left axis) and the share of each month's possible hours classified as Lull, (%, right axis), by calendar month, aggregated across April 2022 to August 2026.*

A Kruskal–Wallis test detected substantial differences among calendar months: daily generation differs significantly across calendar months ($H = 1245.3$, $p < 0.001$) with a large effect size ($\varepsilon^2 = 0.770$), and the same holds for daily lull hours ($H = 1032.2$, $p < 0.001$, $\varepsilon^2 = 0.637$). Table 4 reports results of the test. As noted in Section 3.3, this test assumes independent daily observations, an assumption not separately verified here.

Test variable	H statistic	p-value	Effect size (ε^2)
Daily generation (kWh)	1245.3	< 0.001	0.770 (large)
Daily lull hours	1032.2	< 0.001	0.637 (large)

Table 4. *Kruskal–Wallis test results for daily generation and daily lull hours across the twelve calendar months.*

Figures 4 and 5 show the daily distributions that these tests compare, by calendar month. Both variables are highly skewed and bounded rather than approximately normal: daily generation is zero for most of the year and clusters near zero from January through March, while daily lull hours saturate at or near the 24-hour ceiling in those same months and again from October through December. These figures illustrate the same June-through-September window of consistently elevated generation and low lull hours already evident in Figure 3.

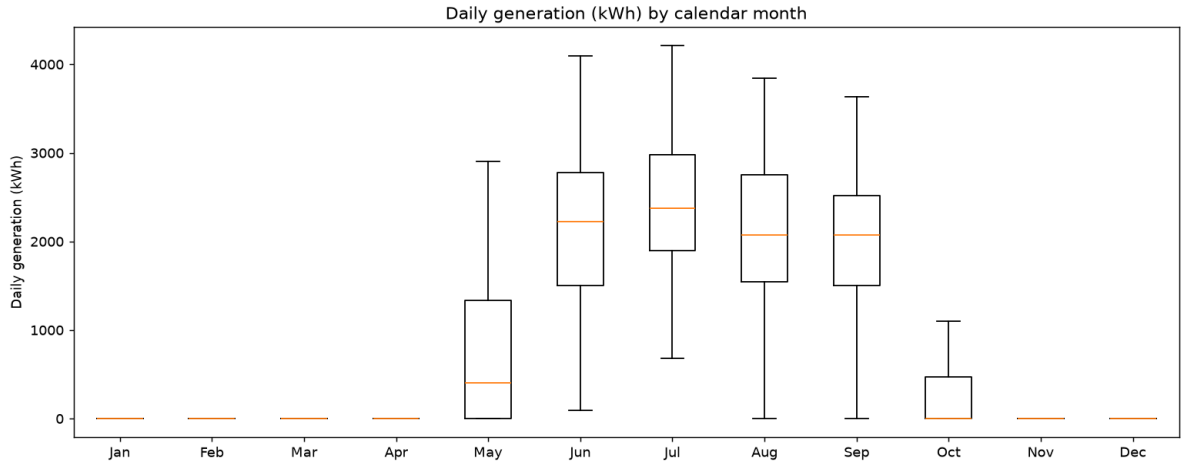

Figure 4. *Distribution of daily generation (kWh) by calendar month, shown as boxplots (median, interquartile range, and range) across all years in the dataset.*

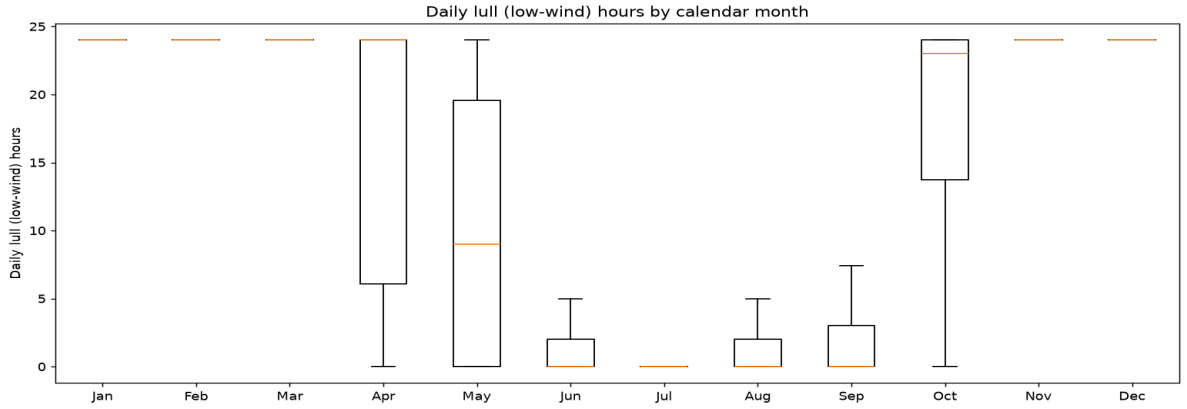

Figure 5. *Distribution of daily lull hours by calendar month, shown as boxplots across all years in the dataset.*

4.3 Reliability: MTBF, MTTR, and Weibull Fit

Breakdown downtime occurred as 29 discrete events across the dataset, with a mean repair duration (MTTR) of 1.28 days (during which an average of 8.14 Breakdown-hours was logged per event), and a mean time between failures (MTBF) of 52.9 days (95% bootstrap CI: 22.0 to 91.2 days). The median gap of 11.5 days is far shorter than the mean, already suggesting an uneven, clustered pattern rather than evenly spaced failures. Grid-Down downtime occurred far more frequently, as 92 discrete events across the dataset, with a mean MTTR of 2.04 days (during which an average of 24.59 Grid-Down-hours was logged per event), and a mean MTBF of 16.6 days (95% CI: 11.5 to 22.7 days; median 8.0 days). Table 5 summarizes these figures together with the fitted Weibull shape and scale parameters for the gaps between events, including bootstrap 95% confidence intervals on both MTBF and the Weibull shape parameter.

Category	N events	MTTR (days)	MTBF (days)	MTBF 95% CI	MTBF median	Weibull k	Weibull k 95% CI	Weibull λ (days)
Breakdown	29	1.28	52.9	22.0 – 91.2	11.5	0.60	0.53 – 0.80	32.7
Grid-Down	92	2.04	16.6	11.5 – 22.7	8.0	0.86	0.76 – 1.06	15.0

***Table 5**. Reliability metrics for Breakdown and Grid-Down downtime events, with bootstrap 95% confidence intervals.*

Both categories returned a Weibull shape point estimate below one (Breakdown: k = 0.60; Grid-Down: k = 0.86), indicating an uneven distribution of inter-event gaps, with short intervals occurring alongside occasional long gaps. This pattern is consistent with, but does not by itself establish, temporal clustering. The strength of this evidence differs between categories: Breakdown's 95% confidence interval (0.53 to 0.80) lies entirely below one, while Grid-Down's confidence interval (0.76 to 1.06) extends slightly above one, meaning the Weibull evidence alone cannot fully rule out a value consistent with random timing for that category. Figures 6 and 7 show these fitted Weibull curves overlaid on the observed gap histograms for Breakdown and Grid-Down, respectively.

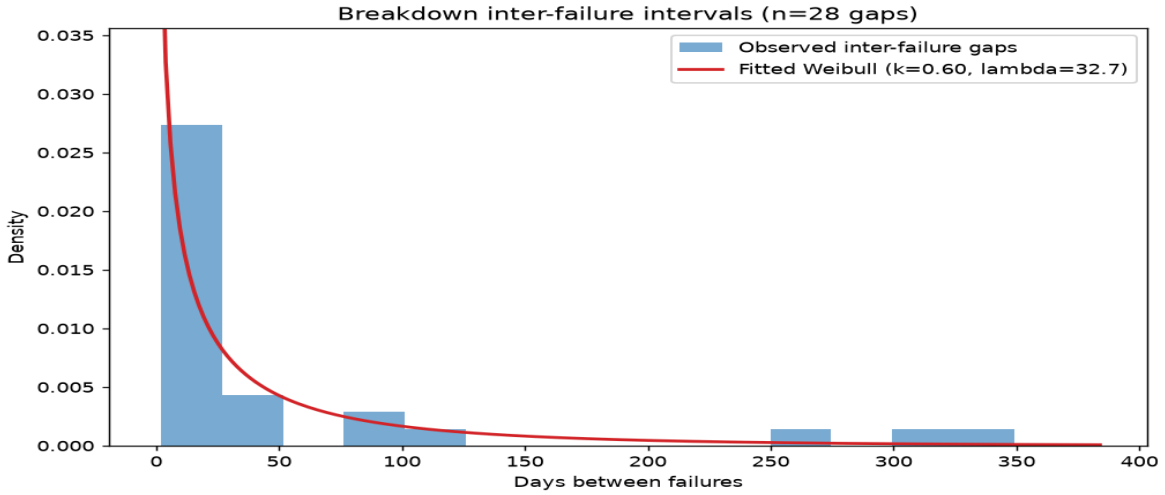

***Figure 6**. Observed Breakdown inter-failure gaps (n = 28) with the fitted Weibull curve overlaid. The y-axis is capped to the histogram's range; since the shape parameter is below 1, the curve's density diverges near zero and extends beyond the top of the plot in that region.*

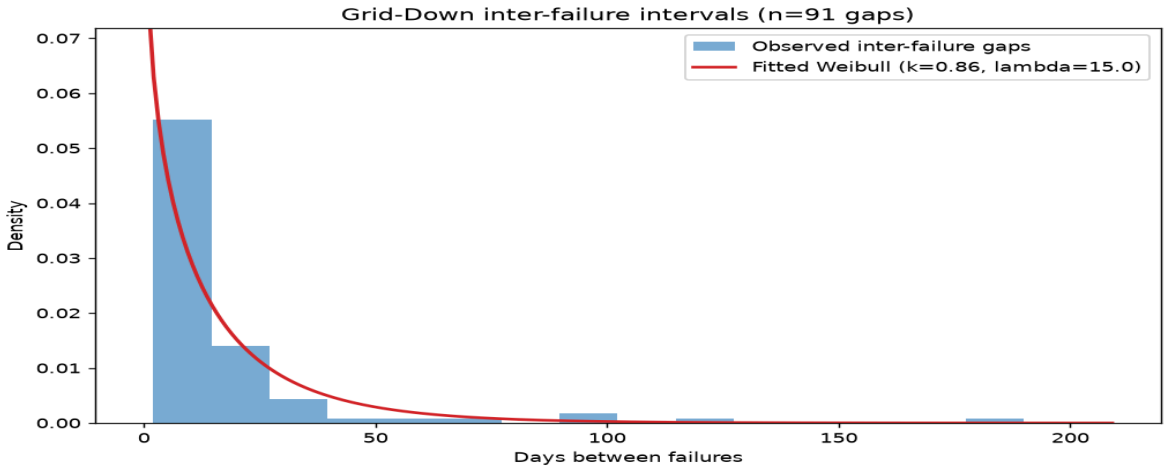

***Figure 7**. Observed Grid-Down inter-failure gaps (n = 91) with the fitted Weibull curve overlaid, on the same basis as Figure 6.*

4.4 Outage Clustering: Kolmogorov–Smirnov Test

As a complementary check, a Kolmogorov–Smirnov (KS) test against an exponential (Poisson-process) null hypothesis, with a Monte Carlo-corrected p-value to account for the rate parameter being estimated from the same data, was run on the same inter-failure gaps for both categories. Table 6 reports the resulting test statistics, and Figures 8 and 9 show the corresponding gap histograms, for Breakdown and Grid-Down respectively, against the fitted exponential null.

Category	N gaps	KS statistic	Asymptotic p	Monte Carlo p	CV
Breakdown	28	0.360	0.0010	0.0005	1.79
Grid-Down	91	0.197	0.0014	0.0005	1.66

***Table 6**. Kolmogorov–Smirnov test of inter-failure gaps against an exponential (Poisson-process) null hypothesis, with an asymptotic and a Monte Carlo-corrected p-value.*

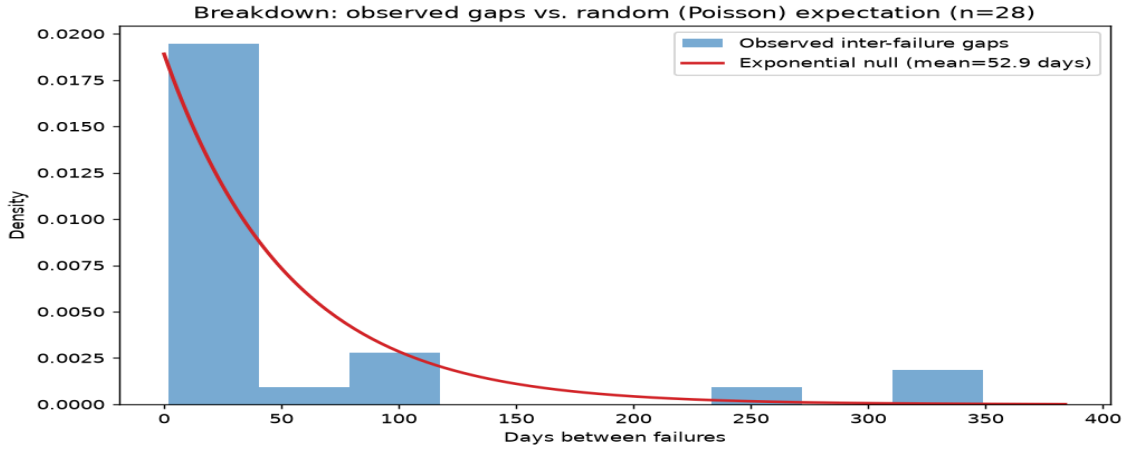

***Figure 8**. Histogram of the observed gaps, in days, between successive Breakdown downtime events (n = 28 gaps, extracted as described in Section 3.3), compared with the exponential distribution expected under a random, constant-rate (homogeneous Poisson) process.*

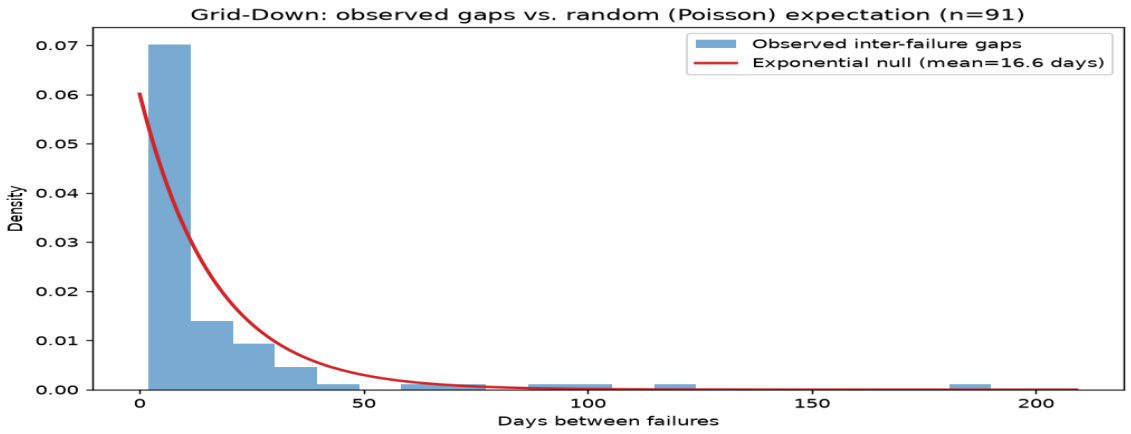

***Figure 9**. Histogram of the observed gaps, in days, between successive Grid-Down downtime events (n = 91 gaps, extracted as described in Section 3.3), compared with the exponential distribution expected under a random, constant-rate (homogeneous Poisson) process.*

For both categories, the Monte Carlo-corrected p-value is 0.0005, so the null hypothesis of random (Poisson) timing is rejected even after correcting for the estimated rate parameter,

and the CV exceeds 1 in both cases (Breakdown: 1.79; Grid-Down: 1.66), consistent with over-dispersion relative to a random process. Figure 8 shows this concretely for Breakdown: with only 28 observed gaps, the bins are coarse, but the histogram shows a dense concentration of short gaps under roughly 40 days, a near-absence of gaps in the 40 to 80 day range, and then a small number of isolated long gaps beyond 240 days, a pattern of tight bursts separated by long quiet stretches that is more consistent with clustering than with the smooth exponential decay a random process would produce. This numerical excess of short gaps is visible directly in Figure 9: the observed histogram bars concentrate at short gap lengths above what the fitted exponential curve predicts. Thus, two complementary statistical methods, a Weibull shape parameter with a confidence interval mostly or entirely below one, and a Monte Carlo-corrected Kolmogorov–Smirnov rejection of the Poisson null, converge on evidence that downtime events at this site are not consistent with a homogeneous, constant-rate random process. Section 5 discusses how this finding should be interpreted.

4.5 Cause of Loss versus Season

Table 7 and Figure 10 cross-tabulate each remarked day's classified cause category against season.

Cause category	Winter	Summer	Monsoon	Post-Monsoon
Grid infrastructure	4	17	15	4
Low wind	86	74	11	55
Maintenance / rework	3	0	2	1
Mechanical / electrical fault	0	8	26	0
Other / unclassified	4	3	17	2
Weather (rain / lightning)	0	29	0	0

***Table** 7. Contingency table of classified cause of loss by season, for the 361 days with a populated remark.*

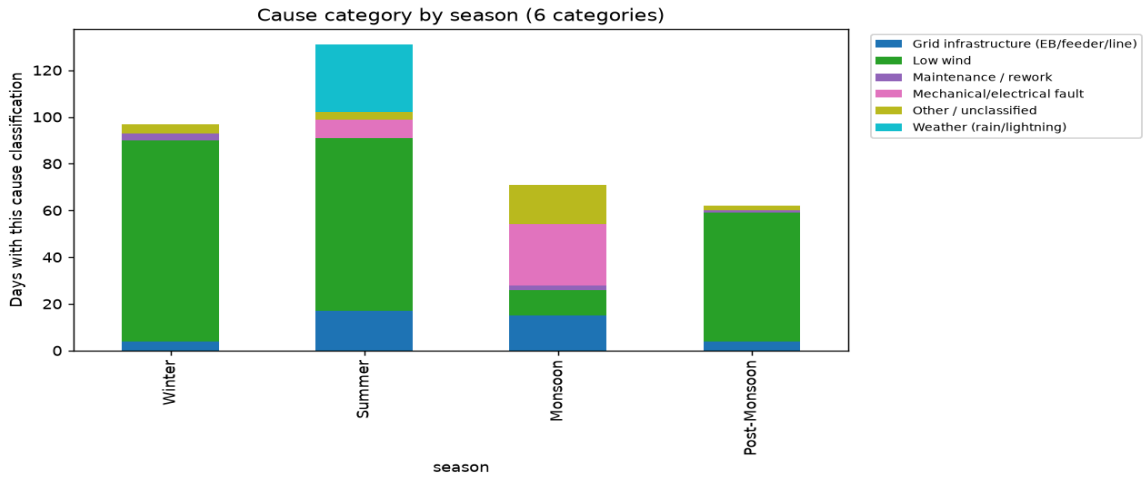

***Figure 10**. Number of remarked days classified into each cause-of-loss category, stacked by season, for the 361 days with a populated remark.*

Table 7 shows that Low Wind is the most frequently classified cause in Winter and Post-Monsoon, while Weather and Mechanical/Electrical Fault causes concentrate in Summer

and Monsoon respectively. Figure 10 shows this pattern visually. Low Wind dominates three of the four seasonal bars: Winter, Summer, and Post-Monsoon. Monsoon is the exception, led instead by Mechanical/Electrical Fault. The six causes do not appear in roughly similar proportions across all four seasons. As noted in Section 3.2, this pattern describes only the 46.5% of Grid-Down and Breakdown hours that occur on a remarked day; the majority of lost hours have no recorded cause and are not represented in Table 7 or Figure 10 at all.

A chi-square test of independence was applied to Table 7. However, only 75.0% of the table's expected cell counts meet the conventional adequacy threshold of five, short of the 80% threshold usually recommended for the chi-square approximation to be considered reliable (Table 7 shows several cells with counts of zero or one). The merged five-category scheme described in Section 3.3.4 raises expected-cell adequacy to 90.0%, and the permutation test on the same merged table provides an independent, assumption-free check. Table 8 reports all three results together: the original chi-square test, the merged-category chi-square test, and the permutation test. The original test found a moderate-to-strong association between cause of loss and season, and this association survives with a similar effect size on the merged scheme. The permutation test likewise rejects the null hypothesis of independence ($p = 0.0005$, the smallest value the 2,000-shuffle resolution can report, since no shuffle produced a statistic as extreme as the one observed).

Test	Statistic	Chi-square p	Permutation p	Cramer's V	Expected cells ≥ 5
6-category (original)	$\chi^2 = 217.34$, df = 15	< 0.001	n/a	0.448	75.0%
Merged 5-category	$\chi^2 = 206.17$, df = 12	< 0.001	n/a	0.436	90.0%
Permutation test (merged, 2,000 shuffles)	n/a	n/a	0.0005	n/a	n/a

***Table 8**. Chi-square test of cause versus season, original and merged category schemes, with a permutation-test cross-check.*

4.6 Economic Sensitivity of Loss Causes

Table 9 ranks the classified cause categories by estimated annual generation opportunity loss and the corresponding estimated revenue, using the site's confirmed tariff of ₹3.39 per kWh and month-specific generation rates as described in Section 3.3, rather than a single site-wide average. Revenue figures are computed from unrounded intermediate kWh values rather than from the rounded kWh/year column shown, so multiplying the displayed kWh/year by the tariff may not reproduce the displayed revenue exactly.

The annual figures scale each category's total estimated loss over the full study period by the number of years the dataset spans, calculated as the date range in days divided by 365.25 (approximately 4.42 years), rather than assuming a fixed 365-day year. The "Days" column reports the number of days assigned to each category: all days without a remark for the unattributed row, and all remarked days classified under that cause for the other rows. It does not indicate the number of days on which Grid-Down or Breakdown hours actually occurred. This refinement changes the ranking among documented causes from what a blanket average would suggest. Weather-related losses are concentrated in Summer, a comparatively low-generation season, so their estimated value falls below that of Grid infrastructure losses once month-specific rates are applied, reversing the order a single average rate would have

produced. The single largest category by a wide margin remains downtime with no remark recorded at all: 1,335.4 hours across 1,253 days, more than the combined total of every specifically documented cause.

Cause category	Lost hours	Days classified	Est. opportunity loss/year (kWh/year)	Est. revenue loss/year (₹/year)
Unattributed (no documented cause)	1,335.4	1,253	20,630	₹69,935
Grid infrastructure	381.5	40	5,786	₹19,615
Weather (rain / lightning)	584.3	29	4,685	₹15,883
Other / unclassified	100.7	26	2,006	₹6,801
Mechanical / electrical fault	83.2	34	1,662	₹5,633
Low wind	9.1	226	198	₹673
Maintenance / rework	4.0	6	87	₹296

Table 9*. Estimated annual generation opportunity loss and revenue loss by classified cause of loss, ranked using month-specific generation rates and the site's confirmed ₹3.39/kWh tariff.*

5. Discussion

The results in Section 4 address four distinct questions about this turbine's four-and-a-half-year operating record, which together support an overall assessment of where and how it lost potential generation.

- **Time allocation**: only 38.3% of possible time was spent generating (Table 2), but the majority of the remaining time (54.5%) was Lull, an absent wind resource rather than an operational failure, while Grid-Down and Breakdown together account for a comparatively small 6.45% of total time.
- **Seasonality**: the near-total January through March generation cessation is both statistically large ($\varepsilon^2 = 0.770$ for generation) and supported by published wind-resource literature for this region (Bastin et al., 2023; George & Banerjee, 2009), consistent with an expected regional wind-resource pattern rather than an unexplained site-specific anomaly (Figures 4 and 5 show the underlying monthly distributions).
- **Failure timing**: both Breakdown and Grid-Down downtime show inter-event gaps inconsistent with a homogeneous Poisson process, supported by two complementary statistical methods (a Weibull shape parameter and a Monte Carlo-corrected Kolmogorov–Smirnov test; Figures 6–9), though this evidence is stronger for Breakdown than for Grid-Down, and is confounded by seasonality rather than proof of true event dependence.
- **Cause attribution**: a chi-square association between cause of loss and season survives two robustness checks (Table 8), but the largest single category of estimated generation opportunity loss is downtime with no documented cause at all (Figure 2, Table 9), covering the majority of actual lost hours. Among the documented causes, applying month-specific rather than site-wide generation rates reverses the ranking of the two largest documented causes by hours lost: Weather-related downtime totals more hours than Grid infrastructure downtime, but is worth less in estimated annual revenue, because Weather losses

concentrate in Summer, a comparatively low-generation season (Table 9). This means the highest-leverage intervention this dataset points toward is improved downtime logging, rather than any single weather or equipment fix. The association itself has a plausible explanation. Low Wind dominates the classified causes in Winter and Post-Monsoon, when the turbine is largely idle for lack of resources. Weather and Mechanical/Electrical Fault causes concentrate in Summer and Monsoon instead, when the turbine is actively generating, and therefore, exposed to both storm damage and operational wear. A related sign of the same documentation gap is that Maintenance hours total only 57 across the entire approximately 4.42-year observation period. This is strikingly low, and some maintenance activity may be being recorded under Breakdown, or not recorded at all, rather than genuinely reflecting so little scheduled servicing.

Together, these findings show that a wind turbine's routine daily reading log, normally kept only to confirm that generation occurred, can support a genuinely quantitative diagnosis of where generation is lost. This is only true if the loss categories are tested statistically, rather than simply summed and charted. The clustering finding illustrates why this distinction matters. A simple monthly bar chart of Grid-Down hours would look the same whether those hours came from many small, unrelated incidents or from a few severe, connected events. The statistical tests in this study point toward the latter, with the seasonal-confounding caveat noted in Section 3.3.3 still applying. For this reason, the clustering result is best read as raising the possibility of a common-cause vulnerability worth investigating further, such as repeated exposure to a shared grid-side vulnerability, rather than as proof that one exists. Confirming it would require the more advanced seasonal modeling identified as future work in Section 8, which the tests used here cannot provide.

In reliability engineering, a Weibull shape parameter near or above one is treated as consistent with random failures, calling for standard fixed-interval maintenance. A shape parameter distinctly below one, as found here for Breakdown ($k = 0.60$), is instead associated with common-cause or bursty failure behavior, and calls for root-cause investigation into what is driving the clustering.

6. Conclusion

This study presents a reproducible framework for decomposing a wind turbine's daily operating log into five time categories and classifying the documented cause of lost time from free-text operational remarks. It then statistically tests three questions about the resulting pattern of loss: whether downtime is consistent with random timing, whether its cause is associated with season, and whether generation differs systematically across calendar months, rather than relying on aggregated totals and charts alone. Demonstrated on a single 250 kW turbine's complete four-and-a-half-year operating history, the framework yielded three findings. First, downtime is inconsistent with random, constant-rate timing, supported by two complementary statistical tests, with appropriate caveats about seasonal confounding. Second, the site's near-total January through March generation cessation is consistent with an expected regional wind-resource pattern. Third, the largest single category of estimated generation opportunity loss is downtime for which no cause was ever recorded. For this installation, the analysis suggests that improving the completeness of downtime documentation may be a higher-leverage first intervention than targeting any single documented failure cause, since the undocumented category alone exceeds every specifically identified cause combined.

Although the demonstration is limited to a single turbine, the methodology depends only on the five time categories and free-text remarks already present in a standard daily reading log, and is therefore potentially applicable to other turbines and sites recording comparable fields, subject to validating the category definitions and data quality at that site. The main contribution of this study is the analytical framework itself, not the specific numbers found at this one installation. By transforming a routinely maintained operational record into a statistically-grounded diagnosis of generation opportunity loss, the framework offers a practical tool for prioritizing maintenance, grid-infrastructure, and record-keeping interventions at small and mid-scale wind installations.

7. Limitations

This framework analyzes a single wind turbine at a single site, over the 1,614 days for which a reading log was available. Further validation using additional turbines and sites would strengthen confidence that the statistical patterns found here, particularly the failure-clustering result, generalize beyond this installation. Within this single site, the dataset itself has one notable gap: it contains no wind speed measurements. Without them, the framework cannot construct a power curve, verify how much of the theoretically available wind resource was actually captured during Run hours, or distinguish a weaker wind year from genuine equipment performance change, so no claim is made in this study about year-over-year efficiency trends.

Several of the statistical tests used in this study rest on assumptions that this dataset cannot fully verify. The clustering result in Section 4.4 carries the seasonal-confounding caveat already noted in Section 3.3.3 and Section 5: it is evidence against constant-rate random timing, but not, by itself, proof that individual events are causally dependent on one another. The Kruskal–Wallis and Kolmogorov–Smirnov tests used in this study also assume independent observations. This assumption is strained by serial correlation in daily weather and operating conditions at a single site across multiple years. It is not separately corrected for here, for example by a block bootstrap or time-stratified test. Also, the daily-resolution event definition used for the reliability analysis (Section 3.3) does not capture within-day start and stop times, which can bias MTTR and MTBF estimates at the margins. Sensitivity of these estimates to alternate event definitions, such as an hour-based activity threshold, is not tested in this study.

The cause classification in Sections 4.5 and 4.6 carries two distinct limitations that compound one another. The classifier itself is a keyword-based rule set rather than a trained or validated model. Independently of the classifier’s accuracy, the remark-coverage gap noted in Section 3.2 means the cause-attribution results characterize a minority of actual lost hours regardless of how accurate the classification is. A further artifact of this day-level attribution is visible in the Low Wind cause category itself: since insufficient wind cannot, by definition, cause Grid-Down or Breakdown hours (that is what the Lull category already captures separately), the negligible 9.1 hours attributed to it in Table 9 most likely reflect incidental Grid-Down or Breakdown minutes falling on days whose remark actually explains the day's Lull hours, rather than a genuine cause of downtime.

A related gap is that this study cannot distinguish Grid-Down time caused by grid curtailment from Grid-Down time caused by faults or outages, since the site’s remarks do not reliably

separate the two. Therefore, treating all Grid-Down time as equally recoverable may overstate the opportunity in Table 9 to the extent curtailment is present.

The economic sensitivity analysis (Section 4.6) improves on a single site-wide average by using month-specific generation rates and the site's confirmed contracted tariff, but still uses a proxy generation rate rather than the wind conditions actually present during any specific downed hour, so the resulting figure remains an estimate rather than an audited loss amount. Finally, Maintenance hours (57 total across the approximately 4.42-year observation period) are strikingly low for this timeframe. This may reflect genuinely infrequent scheduled maintenance, or may indicate that some maintenance activity is logged under a different category, such as Breakdown, or not logged at all, a possibility this study cannot distinguish from the data alone.

8. Future Work

Future work will focus on extending this framework into a more general-purpose analysis tool for operational wind turbine data, and on addressing the statistical confounding identified in Section 7. Other possible extensions include:

- Validation using additional turbines and sites, to test whether the failure-clustering and cause-season association results generalize beyond this installation.
- Modeling failure timing with a non-homogeneous Poisson process or a Cox-type model with an explicit seasonal baseline, to separate genuine event-to-event dependence from a seasonally varying failure rate, which the tests in this study cannot fully disentangle.
- Applying classical renewal-process diagnostics, such as the Laplace trend test or the Lewis-Robinson test, as a complementary check on whether the failure rate itself is trending over time, alongside the Weibull and Kolmogorov–Smirnov evidence reported.
- Testing the sensitivity of MTBF, MTTR, and the Weibull shape parameter to alternate event definitions, such as an hour-based activity threshold instead of a daily one.
- Applying a block-bootstrap or time-stratified variant of the Kruskal–Wallis and Kolmogorov–Smirnov tests, to relax the independence assumption these tests otherwise place on serially correlated daily observations from the same site.
- Quantifying, as a value-of-information question, whether the labor cost of more complete remark-logging, against the ₹69,935/year currently unattributed, is economically worthwhile.
- Hand-labeling a sample of remarks to measure the keyword classifier's precision and recall against a ground-truth set, potentially using a large language model (LLM) for an initial labeling pass followed by human review, following recent demonstrations of LLM-based classification and semantic analysis specifically on wind turbine maintenance logs (Malyi et al., 2025, 2026).
- Incorporating wind speed data, from the site if available or from coarse-resolution reanalysis data (e.g., ERA5-Land) otherwise, to construct a power curve, separate genuine equipment change from weaker wind years, and independently check the Lull classification without disclosing exact coordinates.
- Separating Grid-Down time attributable to grid curtailment from Grid-Down time attributable to grid faults or outages, if the site can supply that distinction, to refine the revenue estimates in Table 9 accordingly.

- Testing whether year-to-year differences in loss decomposition, clustering, and cause-season associations track independently measured monsoon strength, using published regional rainfall data (Harini et al., 2026) as an external covariate, rather than relying on this dataset's generation figures alone to characterize monsoon variability.
- Extending the pipeline to a fleet of turbines, to compare reliability and loss patterns across sites and identify systematically underperforming installations.
- Reporting standard IEC 61400-26-1 availability metrics alongside this study's own categories for comparability, and publishing an example schema or synthetic dataset so the pipeline can be tried without this study's actual data.

9. Acknowledgments

The author is sincerely grateful to Ms. Melo Kamalam, Director of the private limited company that owns the wind turbine analyzed in this study, for generously granting permission to use its operational data for this study. The author also thanks his father, Rajiv Ratnam, whose patient troubleshooting kept the analysis pipeline running through code failures, and whose steady encouragement made the long hours of data wrangling possible.

10. Code Availability

The open-source analysis pipeline developed for this study covering all analyses described in Sections 3 and 4 is publicly available on GitHub (Ratnam, 2026).

11. References

Bastin, J., Haribhaskaran, A., Boopathi, K., Krishnan, B., Vinod Kumar, R., & Reddy Prasad, D. M. (2023). Wind characteristics of Tamil Nadu coast towards development of microgrid: A case study for simulation of small scale hybrid wind and solar energy system. *Ocean Engineering*, 277, Article 114282. https://doi.org/10.1016/j.oceaneng.2023.114282

Dao, C. D., Husayn, I., & Dao, P. B. (2025). Analysis of failures, downtimes and reliability of wind turbines from SCADA data. *Wind Energy*, 28(12), Article e70073. https://doi.org/10.1002/we.70073

George, M., & Banerjee, R. (2009). Analysis of impacts of wind integration in the Tamil Nadu grid. *Energy Policy*, 37(9), 3693–3700. https://doi.org/10.1016/j.enpol.2009.04.045

Harini, S., Sharma, D., Patil, Y., Chopra, G., Tandon, S., Goswami, B. N., & Sujith, R. I. (2026). Prediction and predictability of the wet-season rainfall over Southeast India. *arXiv*. https://doi.org/10.48550/arXiv.2605.01326

International Electrotechnical Commission [IEC]. (2019). *IEC 61400-26-1:2019 – Wind energy generation systems – Part 26-1: Availability for wind energy generation systems.* https://webstore.iec.ch/en/publication/62548

International Renewable Energy Agency [IRENA]. (n.d.). *Wind energy.* IRENA. https://www.irena.org/Energy-Transition/Technology/Wind-energy

Malyi, M., Shek, J., & Biscaya, A. (2025). Exploratory semantic reliability analysis of wind turbine maintenance logs using large language models. *arXiv*. https://doi.org/10.48550/arXiv.2509.22366

Malyi, M., Shek, J., McDonald, A., & Biscaya, A. (2026). A comparative benchmark of large language models for labelling wind turbine maintenance logs. *IET Renewable Power Generation*, 20(1), Article e70212. https://doi.org/10.1049/rpg2.70212

McDonald, J. H. (2014). *Handbook of biological statistics* (3rd ed.). Sparky House Publishing. http://www.biostathandbook.com/

Ratnam, A. R. (2026). *wind-turbine-loss-analysis* [Computer software]. GitHub. https://github.com/adityaratnam09/wind-turbine-loss-analysis

Truong, B.-C., Mphekgwana, P., & Pal, N. (2025). A revisit to maximum likelihood estimation of Weibull model parameters. *arXiv*. https://doi.org/10.48550/arXiv.2501.11604